

P1 1173878

REC'D 26 MAY 2004

WIPO

PCT

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office

May 21, 2004

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: 60/460,526

FILING DATE: April 03, 2003

RELATED PCT APPLICATION NUMBER: PCT/US04/10582

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS



M. Tarver
M. TARVER
Certifying Officer

**PRIORITY
DOCUMENT**

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

BEST AVAILABLE COPY

04-07-03

60463526 040304


04/03/03
 35682 U.S. PTO

Approved for use through 10/31/2002. OMB 0651-0032
 U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
 Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No. EV174834883US

INVENTOR(S)			
Given Name (first and middle (if any))	Family Name or Surname	Residence (City and either State or Foreign Country)	
Iris Tsvika Shmuel	Alroy Greener Tuvia	Nes Ziona, Israel Nes Ziona, Israel Netanya, Israel	
<input checked="" type="checkbox"/> Additional inventors are being named on the <u>1</u> separately numbered sheets attached hereto			
TITLE OF THE INVENTION (500 characters max)			
POSH AND ASSOCIATED PROTEINS			
Direct all correspondence to:		CORRESPONDENCE ADDRESS	
<input checked="" type="checkbox"/> Customer Number		 * 2 8 1 2 0 * Customer Number Bar Code	
<input checked="" type="checkbox"/> Firm or Individual Name		Anita Varma ROPES & GRAY	
Address			
City	State	Zip	
Country	Telephone	Fax	
ENCLOSED APPLICATION PARTS (check all that apply)			
<input checked="" type="checkbox"/> Specification Number of Pages	141	<input type="checkbox"/> CD(s), Number	
<input checked="" type="checkbox"/> Drawing(s) Number of Sheets	29	<input type="checkbox"/> Other:	
<input type="checkbox"/> Application Data Sheet. See 37 CFR 1.76			
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT			
<input checked="" type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27.		FILING FEE AMOUNT (\$)	
<input type="checkbox"/> A check or money order is enclosed to cover the filing fees			
<input checked="" type="checkbox"/> The Commissioner is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number:		18-1945	80.00
<input type="checkbox"/> Payment by credit card. Form PTO-2038 is attached.			
The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.			
<input checked="" type="checkbox"/> No		<input type="checkbox"/> Yes, the name of the U.S. Government agency and the Government contract number are:	

Respectfully submitted,

SIGNATURE

TYPED OR
PRINTED NAME

TELEPHONE

Robert A. Mazzerese

(617) 951-7796

Date April 3, 2003

REGISTRATION NO.
(if appropriate)

42,852

Docket Number:

PROL-P67-024

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

I hereby certify that this correspondence is being deposited with the U.S. Postal Service as Express Mail, Airbill No. EV174834883US, in an envelope addressed to: Box Provisional Patent Application, Commissioner for Patents, Washington, DC 20231, on the date shown below.

Dated: 4/3/03

Signature: Maura A. Gallagher (Maura A. Gallagher)

60460526 . 040303

PROVISIONAL APPLICATION COVER SHEET **Additional Page**

PTO/SB/16 (10-01)

Approved for use through 10/31/2002. OMB 0651-0032
 U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Docket Number		PROL-P67-024
INVENTOR(S)/APPLICANT(S)		
Given Name (first and middle (if any))	Family or Surname	Residence (City and either State or Foreign Country)
Danny	Ben-Avraham	Zichron Jackov, Israel

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

**FEE TRANSMITTAL
for FY 2003**

Patent fees are subject to annual revision.

☒ Applicant claims small entity status. -See 37 CFR 1.27**TOTAL AMOUNT OF PAYMENT (\$)** 80.00**Complete if Known**

Application Number	Not Yet Assigned
Filing Date	April 3, 2003
First Named Inventor	Iris Alroy
Examiner Name	Not Yet Assigned
Group Art Unit	N/A
Attorney Docket No.	PROL-P67-024

METHOD OF PAYMENT (check all that apply)
☐ Check ☐ Credit Card ☐ Money Order ☐ Other ☐ None
☒ Deposit AccountDeposit Account Number **18-1945**Deposit Account Name **Ropes & Gray**

The Commissioner is hereby authorized to: (check all that apply)

☒ Charge fee(s) indicated below ☒ Credit any overpayments☒ Charge any additional fee(s) during the pendency of this application☐ Charge fee(s) indicated below, except for the filing fee

to the above-identified deposit account.

FEE CALCULATION**1. BASIC FILING FEE**

Large Entity		Small Entity		Fee Description	Fee Paid
Fee Code	Fee (\$)	Fee Code	Fee (\$)		
1001	750	2001	375	Utility filing fee	
1002	330	2002	165	Design filing fee	
1003	520	2003	260	Plant filing fee	
1004	750	2004	375	Reissue filing fee	
1005	160	2005	80	Provisional filing fee	80.00
SUBTOTAL (1)					80.00

2. EXTRA CLAIM FEES FOR UTILITY AND REISSUE

Total Claims	Extra Claims	Fee from below	Fee Paid
Independent Claims	.. =	x	=
Multiple Dependent	.. =	x	=

Large Entity		Small Entity		Fee Description	Fee Paid
Fee Code	Fee (\$)	Fee Code	Fee (\$)		
1202	18	2202	9	Claims in excess of 20	
1201	84	2201	42	Independent claims in excess of 3	
1203	280	2203	140	Multiple dependent claim, if not paid	
1204	84	2204	42	** Reissue independent claims over original patent	
1205	18	2205	9	** Reissue claims in excess of 20 and over original patent	

SUBTOTAL (2) (\$)

0.00

**or number previously paid, if greater; For Reissues, see above

FEE CALCULATION (continued)**3. ADDITIONAL FEES**

Large Entity		Small Entity		Fee Description	Fee Paid
Fee Code	Fee (\$)	Fee Code	Fee (\$)		
1051	130	2051	65	Surcharge - late filing fee or oath	
1052	50	2052	25	Surcharge - late provisional filing fee or cover sheet	
1053	130	1053	130	Non-English specification	
1812	2,520	1812	2,520	For filing a request for <i>ex parte</i> reexamination	
1804	920*	1804	920*	Requesting publication of SIR prior to Examiner action	
1805	1,840*	1805	1,840*	Requesting publication of SIR after Examiner action	
1251	110	2251	55	Extension for reply within first month	
1252	410	2252	205	Extension for reply within second month	
1253	930	2253	465	Extension for reply within third month	
1254	1,450	2254	725	Extension for reply within fourth month	
1255	1,970	2255	985	Extension for reply within fifth month	
1401	320	2401	160	Notice of Appeal	
1402	320	2402	160	Filing a brief in support of an appeal	
1403	280	2403	140	Request for oral hearing	
1451	1,510	1451	1,510	Petition to institute a public use proceeding	
1452	110	2452	55	Petition to revive - unavoidable	
1453	1,300	2453	650	Petition to revive - unintentional	
1501	1,300	2501	650	Utility issue fee (or reissue)	
1502	470	2502	235	Design issue fee	
1503	630	2503	315	Plant issue fee	
1460	130	1460	130	Petitions to the Commissioner	
1807	50	1807	50	Processing fee under 37 CFR 1.17(q)	
1806	180	1806	180	Submission of Information Disclosure Stmt	
8021	40	8021	40	Recording each patent assignment per property (times number of properties)	
1809	750	2809	375	Filing a submission after final rejection (37 CFR 1.129(a))	
1810	750	2810	375	For each additional invention to be examined (37CFR 1.129(b))	
1801	750	2801	375	Request for Continued Examination (RCE)	
1802	800	1802	800	Request for expedited examination of a design application	

Other fee (specify)

*Reduced by Basic Filing Fee Paid

SUBTOTAL (3) (\$)

0.00

SUBMITTED BYName (Print/Type) **Robert A. Mazzarese**

Registration No. (Attorney/Agent)

42,852

Complete (if applicable)

Telephone (617) 951-7804

Signature

Date

April 3, 2003

I hereby certify that this correspondence is being deposited with the U.S. Postal Service as Express Mail, Airbill No. EV174834883US, in an envelope addressed to: Box Provisional Patent Application, Commissioner for Patents, Washington, DC 20231, on the date shown below.

Dated: 4/3/03Signature: Maura A. Gallagher (Maura A. Gallagher)

POSH AND ASSOCIATED PROTEINS

BACKGROUND

Potential drug target validation involves determining whether a DNA, RNA
5 or protein molecule is implicated in a disease process and is therefore a suitable
target for development of new therapeutic drugs. Drug discovery, the process by
which bioactive compounds are identified and characterized, is a critical step in the
development of new treatments for human diseases. The landscape of drug
discovery has changed dramatically due to the genomics revolution. DNA and
10 protein sequences are yielding a host of new drug targets and an enormous amount
of associated information.

The identification of genes and proteins involved in various disease states or
key biological processes, such as inflammation and immune response, is a vital part
of the drug design process. Many diseases and disorders could be treated or
15 prevented by decreasing the expression of one or more genes involved in the
molecular etiology of the condition if the appropriate molecular target could be
identified and appropriate antagonists developed. For example, cancer, in which one
or more cellular oncogenes become activated and result in the unchecked
progression of cell cycle processes, could be treated by antagonizing appropriate cell
20 cycle control genes. Furthermore many human genetic diseases, such as
Huntington's disease, and certain prion conditions, which are influenced by both
genetic and epigenetic factors, result from the inappropriate activity of a polypeptide
as opposed to the complete loss of its function. Accordingly, antagonizing the
aberrant function of such mutant genes would provide a means of treatment.
25 Additionally, infectious diseases such as HIV have been successfully treated with
molecular antagonists targeted to specific essential retroviral proteins such as HIV
protease or reverse transcriptase. Drug therapy strategies for treating such diseases
and disorders have frequently employed molecular antagonists which target the
polypeptide product of the disease gene(s). However the discovery of relevant gene
30 or protein targets is often difficult and time consuming.

One area of particular interest is the identification of host genes and proteins
that are co-opted by viruses during the viral life cycle. The serious and incurable

nature of many viral diseases, coupled with the high rate of mutations found in many viruses, makes the identification of antiviral agents a high priority for the improvement of world health. Genes and proteins involved in a viral life cycle are also appealing as a subject for investigation because such genes and proteins will typically have additional activities in the host cell and may play a role in other non-viral disease states.

Viral maturation involves the proteolytic processing of the Gag proteins and the activity of various host proteins. It is believed that cellular machineries for exo/endocytosis and for ubiquitin conjugation may be involved in the maturation. In particular, the assembly, budding and subsequent release of reovirus, RNA viruses and envelope viruses, such as various retroviruses, rhabdoviruses, lentiviruses, and filoviruses may involve the Gag polyprotein. After its synthesis, Gag is targeted to the plasma membrane where it induces budding of nascent virus particles.

The role of ubiquitin in virus assembly was suggested by Dunigan et al. (1988, Virology 165, 310, Meyers et al. 1991, Virology 180, 602), who observed that mature virus particles were enriched in unconjugated ubiquitin. More recently, it was shown that proteasome inhibitors suppress the release of HIV-1, HIV-2 and virus-like particles derived from SIV and RSV Gag. Also, inhibitors affect Gag processing and maturation into infectious particles (Schubert et al 2000, PNAS 97, 13057, Harty et al. 2000, PNAS 97, 13871, Strack et al. 2000, PNAS 97, 13063, Patnaik et al. 2000, PNAS 97, 13069).

It is well known in the art that ubiquitin-mediated proteolysis is the major pathway for the selective, controlled degradation of intracellular proteins in eukaryotic cells. Ubiquitin modification of a variety of protein targets within the cell appears to be important in a number of basic cellular functions such as regulation of gene expression, regulation of the cell-cycle, modification of cell surface receptors, biogenesis of ribosomes, and DNA repair. One major function of the ubiquitin-mediated system is to control the half-lives of cellular proteins. The half-life of different proteins can range from a few minutes to several days, and can vary considerably depending on the cell-type, nutritional and environmental conditions, as well as the stage of the cell-cycle.

Targeted proteins undergoing selective degradation, presumably through the actions of a ubiquitin-dependent proteasome, are covalently tagged with ubiquitin through the formation of an isopeptide bond between the C-terminal glycyl residue of ubiquitin and a specific lysyl residue in the substrate protein. This process is catalyzed by a ubiquitin-activating enzyme (E1) and a ubiquitin-conjugating enzyme (E2), and in some instances may also require auxiliary substrate recognition proteins (E3s). Following the linkage of the first ubiquitin chain, additional molecules of ubiquitin may be attached to lysine side chains of the previously conjugated moiety to form branched multi-ubiquitin chains.

10 The conjugation of ubiquitin to protein substrates is a multi-step process. In an initial ATP requiring step, a thioester is formed between the C-terminus of ubiquitin and an internal cysteine residue of an E1 enzyme. Activated ubiquitin may then be transferred to a specific cysteine on one of several E2 enzymes. Finally, these E2 enzymes donate ubiquitin to protein substrates, typically with the assistance of an E3 protein, also known as a ubiquitin enzyme. In certain instances, substrates are recognized directly by the ubiquitin-conjugated E2 enzyme.

It is also known that the ubiquitin system plays a role in a wide range of cellular processes including cell cycle progression, apoptosis, and turnover of many membrane receptors. In viral infections, the ubiquitin system is involved not only with assembly, budding and release, but also with repression of host proteins such as p53, which may lead to a viral-induced neoplasm. The HIV Vpu protein interacts with an E3 protein that regulates I κ B degradation, and is thought to promote apoptosis of infected cells by indirectly inhibiting NF- κ B activity (Bour et al. (2001) J Exp Med 194:1299-311; U.S. Patent No. 5,932,425). The ubiquitin system regulates protein function by both mono-ubiquitination and poly-ubiquitination, and poly-ubiquitination is primarily associated with protein degradation.

20 The vesicular trafficking systems are the major pathways for the distribution of proteins among cell organelles, the plasma membrane and the extracellular medium. The vesicular trafficking systems may be directly or indirectly involved in a variety of disease states. The major vesicle trafficking systems in eukaryotic cells include those systems that are mediated by clathrin-coated vesicles and coatomer-coated vesicles. Clathrin-coated vesicles are generally involved in transport, such as

in the case of receptor mediated endocytosis, between the plasma membrane and the early endosomes, as well as from the trans-Golgi network to endosomes. Coatomer-coated vesicles include coat protein I (COP-I) coated vesicles and COP-II coated vesicles, both of which tend to mediate transport of a variety of molecules between the ER and Golgi cisternae. In each case, a vesicle is formed by budding out from a portion of membrane that is coated with coat proteins, and the vesicle sheds its coat prior to fusing with the target membrane.

Clathrin coats assemble on the cytoplasmic face of a membrane, forming pits that ultimately pinch off to become vesicles. Clathrin itself is composed of two subunits, the clathrin heavy chain and the clathrin light chain, that form the clathrin triskelion. Clathrins associate with a host of other proteins, including the assembly protein, AP180, the adaptor complexes (AP1, AP2, AP3 and AP4), beta-arrestin, arrestin 3, auxilin, epsin, Eps15, v-SNAREs, amphiphysins, dynamin, synaptojanin and endophilin. The adaptor complexes promote clathrin cage formation, and help connect clathrin up to the membrane, membrane proteins, and many of the preceding components. AP1 associates with clathrin coated vesicles derived from the trans-Golgi network and contains γ , $\beta 1$, $\mu 1$ and $\sigma 1$ polypeptide chains. AP2 associates with endocytic clathrin coated vesicles and contains α , $\beta 2$, $\mu 2$, and $\sigma 2$ polypeptides. Interactions between the clathrin complex and other proteins are mediated by a variety of domains found in the complex proteins, such as SH3 (Src homology 3) domains, PH (pleckstrin homology) domains, EH domains and NPF domains. (Marsh et al. (1999) Science 285:215-20; Pearse et al. (2000) Curr Opin Struct Biol 10(2):220-8).

Coatomer-coated vesicle formation is initiated by recruitment of a small GTPase (eg. ARF or SAR) by its cognate guanine nucleotide exchange factor (e.g. SEC12, GEA1, GEA2). The initial complex is recognized by a coat protein complex (COPI or COPII). The coat then grows across the membrane, and various cargo proteins become entrapped in the growing network. The membrane ultimately bulges and becomes a vesicle. The coat proteins stimulate the GTPase activity of the GTPase, and upon hydrolysis of the GTP, the coat proteins are released from the complex, uncoating the vesicle. Other proteins associated with coatomer coated

vesicles include v-SNAREs, Rab GTPases and various receptors that help recruit the appropriate cargo proteins. (Springer et al. (1999) Cell 97:145-48).

It would be beneficial to identify proteins involved in one or more of these processes for use in, among other things, drug screening methods.

5

SUMMARY

In part, the application relates to the ubiquitin ligase, POSH (Plenty Of SH3 domains), and the discovery of novel interactions between POSH and proteins that associate with POSH. By providing novel POSH:POSH-AP interactions, the application provides, in part, methods for modulating a process that POSH participates in by targeting a POSH-AP or the POSH:POSH-AP interaction. Furthermore, by providing novel POSH:POSH-AP interactions, the application provides, in part, methods for modulating a process that a POSH-AP participates in by targeting POSH. As one of skill in the art can readily appreciate, a POSH protein may form multiple different complexes with POSH-APs, depending on the biological context.

In certain embodiments, POSH and POSH-associated proteins (POSH-APs) play a role in viral maturation. Optionally, POSH and POSH-APs act in the assembly or trafficking of complexes that mediate viral release. In one embodiment, POSH polypeptides and POSH-APs may stimulate ubiquitylation of certain proteins or stimulate membrane fusion or both. In certain embodiments, POSH and POSH-APs participate in vesicular trafficking. In certain embodiments, POSH and POSH-APs regulate a Rac signaling pathway. In certain embodiments, POSH and POSH-APs regulate a JNK signaling pathway. In certain embodiments, POSH and POSH-APs regulate NF-kB and NF-kB signaling. In certain embodiments, POSH and POSH-APs participate in autoregulation of POSH polypeptide levels. In certain embodiments, POSH and POSH-APs regulate apoptosis. In certain embodiments, POSH and POSH-APs participate in cellular positioning of the nucleus. In certain embodiments, POSH and POSH-APs participate in attachment of the nuclear membrane to the actin cytoskeleton. In certain embodiments, POSH and POSH-APs participate in cellular responses to homocysteine, including the unfolded protein response (UPR), thromboembolic vascular disease, apoptosis of vascular endothelial

cells, dysregulation of sterol synthesis (e.g. cholesterol synthesis) positioning of the nucleus.

In some aspects, the application provides POSH and POSH-AP nucleic acid sequences and proteins encoded thereby, as well as oligonucleotides derived from the nucleic acid sequences, antibodies directed to the encoded proteins, screening assays to identify agents that modulate POSH, POSH-APs and/or biological events affected by POSH and POSH-APs. In certain aspects the application provides diagnostic methods for detecting cells infected with a virus, preferably an envelope virus, an RNA virus and particularly a retrovirus.

In one aspect, the application provides an isolated nucleic acid comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence of SEQ ID NOs: 1, 3, 4, 6, 8 and/or 10 or a sequence complementary thereto. In a related embodiment, the nucleic acid is at least about 80%, 90%, 95%, or 97-98%, or 100% identical to a sequence corresponding to at least about 12, at least about 15, at least about 25, at least about 40, at least about 100, at least about 300, at least about 500, at least about 1000, or at least about 2500 consecutive nucleotides up to the full length of SEQ ID NO: 1, 3, 4, 6, 8 and/or 10, or a sequence complementary thereto.

In one aspect, the application provides an isolated nucleic acid comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence of SEQ ID NOs: 31-35 or a sequence complementary thereto. In a related embodiment, the nucleic acid is at least about 80%, 90%, 95%, or 97-98%, or 100% identical to a sequence corresponding to at least about 12, at least about 15, at least about 25, consecutive nucleotides up to the full length of SEQ ID NO: 31-35, or a sequence complementary thereto.

In other embodiments, the application provides a nucleic acid comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence of SEQ ID Nos. 1, 3, 4, 6, 8 and/or 10, or a nucleotide sequence that is at least about 80%, 90%, 95%, or 97-98%, or 100% identical to a sequence corresponding to at least about 12, at least about 15, at least about 25, at least about 40, at least about 100, at least about 300, at least about 500, at least about 1000, or at least about 2500 consecutive nucleotides up to the full length of SEQ ID NO: 1, 3, 4, 6, 8 and/or 10,

or a sequence complementary thereto, and a transcriptional regulatory sequence operably linked to the nucleotide sequence to render the nucleotide sequence suitable for use as an expression vector. In another embodiment, the nucleic acid may be included in an expression vector capable of replicating in a prokaryotic or eukaryotic cell. In a related embodiment, the application provides a host cell transfected with the expression vector.

In yet another embodiment, the application provides a substantially pure nucleic acid which hybridizes under stringent conditions to a nucleic acid probe corresponding to at least about 12, at least about 15, at least about 25, or at least about 40 consecutive nucleotides up to the full length of SEQ ID NO:1, 3, 4, 6, 8 and/or 10, or a sequence complementary thereto or up to the full length of the gene of which said sequence is a fragment. The application also provides an antisense oligonucleotide analog which hybridizes under stringent conditions to at least 12, at least 25, or at least 50 consecutive nucleotides up to the full length of SEQ ID NO:1 and/or 3, or a sequence complementary thereto.

In a further embodiment, the application provides a nucleic acid comprising a nucleic acid encoding an amino acid sequence as set forth in any of SEQ ID Nos: 2, 5, 7, 9 or 11, or a nucleic acid complement thereof. In a related embodiment, the encoded amino acid sequence is at least about 80%, 90%, 95%, or 97-98%, or 100% identical to a sequence corresponding to at least about 12, at least about 15, at least about 25, or at least about 40, at least about 100, at least about 200, at least about 300, at least about 400 or at least about 500 consecutive amino acids up to the full length of any of SEQ ID Nos:2, 5, 7, 9 or 11.

In another embodiment, the application provides a probe/primer comprising a substantially purified oligonucleotide, said oligonucleotide containing a region of nucleotide sequence which hybridizes under stringent conditions to at least about 12, at least about 15, at least about 25, or at least about 40 consecutive nucleotides of sense or antisense sequence selected from SEQ ID Nos: 1, 3, 4, 6, 8 and/or 10, or a sequence complementary thereto. In preferred embodiments, the probe selectively hybridizes with a target nucleic acid. In another embodiment, the probe may include a label group attached thereto and able to be detected. The label group may be selected from radioisotopes, fluorescent compounds, enzymes, and enzyme co-

factors. The application further provides arrays of at least about 10, at least about 25, at least about 50, or at least about 100 different probes as described above attached to a solid support.

In another aspect, the application provides polypeptides. In one embodiment, the application pertains to a polypeptide including an amino acid sequence encoded by a nucleic acid comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence of SEQ ID Nos:1, 3, 4, 6, 8 and/or 10, or a sequence complementary thereto, or a fragment comprising at least about 25, or at least about 40 amino acids thereof.

In a preferred embodiment, the POSH polypeptide comprises a sequence that is identical with or homologous to any of SEQ ID Nos: 2, 5, 7, 9 or 11. For instance, a POSH polypeptide preferably has an amino acid sequence at least 60% homologous to a polypeptide represented by any of SEQ ID Nos:2, 5, 7, 9 or 11 and polypeptides with higher sequence homologies of, for example, 80%, 90% or 95% are also contemplated. The POSH polypeptide can comprise a full length protein, such as represented in the sequence listings, or it can comprise a fragment of, for instance, at least 5, 10, 20, 50, 100, 150, 200, 250, 300, 400 or 500 or more amino acids in length.

In another embodiment, the application provides polypeptides comprising a sequence that is at least 80%, 90% or 95% identical with or homologous to any of SEQ ID Nos: 26-30.

In another preferred embodiment, the application features a purified or recombinant polypeptide fragment of a POSH polypeptide, which polypeptide has the ability to modulate, e.g., mimic or antagonize, an activity of a wild-type POSH protein. Preferably, the polypeptide fragment comprises a sequence identical or homologous to an amino acid sequence designated in any of SEQ ID Nos: 2, 5, 7, 9 or 11.

Moreover, as described below, the POSH polypeptide can be either an agonist (e.g. mimics), or alternatively, an antagonist of a biological activity of a naturally occurring form of the protein, e.g., the polypeptide is able to modulate the intrinsic biological activity of a POSH protein or a POSH complex, such as an

enzymatic activity, binding to other cellular components, cellular compartmentalization, membrane reorganization and the like.

The subject proteins can also be provided as chimeric molecules, such as in the form of fusion proteins. For instance, the POSH polypeptide can be provided as
 5 a recombinant fusion protein which includes a second polypeptide portion, e.g., a second polypeptide having an amino acid sequence unrelated (heterologous) to POSH, e.g. the second polypeptide portion is glutathione-S-transferase, e.g. the second polypeptide portion is an enzymatic activity such as alkaline phosphatase, e.g. the second polypeptide portion is an epitope tag, etc.

10 Yet another aspect of the present application concerns an immunogen comprising a POSH polypeptide in an immunogenic preparation, the immunogen being capable of eliciting an immune response specific for the POSH polypeptide; e.g. a humoral response, e.g. an antibody response; e.g. a cellular response. In preferred embodiments, the immunogen comprises an antigenic determinant, e.g. a
 15 unique determinant, from a protein represented by SEQ ID NO:2.

In yet another aspect, this application provides antibodies immunoreactive with one or more POSH polypeptides. In one embodiment, antibodies are specific for an SH3 domain or a RING domain derived from a POSH polypeptide. In a more specific embodiment, the domain is part of an amino acid sequence set forth in SEQ
 20 ID NO:2. In a set of exemplary embodiments, an antibody binds to one or more SH3 domains represented by amino acids 137-192, 199-258, 448-505 and 832-888 of SEQ ID NO:2 and are set forth in any one of SEQ ID Nos: 27-20. In another exemplary embodiment, an antibody binds to a RING domain represented by amino acids 12-52 of SEQ ID NO:2 and is set forth in SEQ ID No: 26. In another
 25 embodiment, the antibodies are immunoreactive with one or more proteins having an amino acid sequence that is at least 80% identical, at least 90% identical or at least 95% identical to an amino acid sequence as set forth in SEQ ID NO:2. In other embodiments, an antibody is immunoreactive with one or more proteins having an amino acid sequence that is 85%, 90%, 95%, 98%, 99% or identical to an amino
 30 acid sequence as set forth in SEQ ID NO:2.

In certain embodiments, the subject POSH nucleic acids will include a transcriptional regulatory sequence, e.g. at least one of a transcriptional promoter or

transcriptional enhancer sequence, which regulatory sequence is operably linked to the POSH sequence. Such regulatory sequences can be used to render the POSH sequence suitable for use as an expression vector.

In yet another aspect, the application provides an assay for screening test compounds for inhibitors, or alternatively, potentiators, of an interaction between a POSH polypeptide and a POSH-AP. An exemplary method includes the steps of (i) combining POSH-AP, a POSH polypeptide, and a test compound, e.g., under conditions wherein, but for the test compound, the POSH polypeptide and POSH-AP are able to interact; and (ii) detecting the formation of a complex which includes the POSH polypeptide and a POSH-AP. A statistically significant change, such as a decrease, in the formation of the complex in the presence of a test compound (relative to what is seen in the absence of the test compound) is indicative of a modulation, e.g., inhibition, of the interaction between the POSH polypeptide and POSH-AP.

In a further embodiment, the application provides an assay for identifying a test compound which inhibits or potentiates the interaction of a POSH polypeptide to a POSH-AP, comprising (a) forming a reaction mixture including POSH polypeptide, a POSH-AP; and a test compound; and detecting binding of said POSH polypeptide to said POSH-AP; wherein a change in the binding of said POSH polypeptide to said POSH-AP in the presence of the test compound, relative to binding in the absence of the test compound, indicates that said test compound potentiates or inhibits binding of said POSH polypeptide to said POSH-AP.

In additional embodiment, the application relates to a method for identifying modulators of protein complexes, comprising (a) forming a reaction mixture comprising a POSH polypeptide, a POSH-AP; and a test compound; (b) contacting the reaction mixture with a test agent, and (c) determining the effect of the test agent for one or more activities. Exemplary activities include a change in the level of the protein complex, a change in the enzymatic activity of the complex, where the reaction mixture is a whole cell, a change in the plasma membrane localization of the complex or a component thereof or a change in the interaction between the POSH polypeptide and the POSH-AP.

An additional embodiment is a screening assay to identify agents that inhibit or potentiate the interaction of a POSH polypeptide and a POSH-AP, comprising providing a two-hybrid assay system including a first fusion protein comprising a POSH polypeptide portion of SEQ ID NO:2, and a second fusion protein comprising a POSH-AP portion, under conditions wherein said two hybrid assay is sensitive to interactions between the POSH polypeptide portion of said first fusion protein and said POSH-AP portion of said second polypeptide; measuring a level of interactions between said fusion proteins in the presence and in the absence of a test agent; and comparing the level of interaction of said fusion proteins, wherein a decrease in the level of interaction is indicative of an agent that will inhibit the interaction between a POSH polypeptide and a POSH-AP.

In additional aspects, the application provides isolated protein complexes including a combination of a POSH polypeptide and at least one POSH-AP. In certain embodiments, a POSH complex is related to clathrin-coated vesicle formation. In a further embodiment, a POSH complex comprises a viral protein, such as Gag. In certain embodiments, a POSH complex relates to a ubiquitin related activity of POSH, as in the case of POSH complexes comprising ubiquitin (e.g., covalent or non-covalent POSH ubiquitin conjugates), an E2, an E1 or a ubiquitination target. In certain embodiments, a POSH complex relates to JNK signaling, as in the case of POSH complexes comprising a Rac, an MLK, an MKK and/or a JNK. In certain embodiments, a POSH complex is part of a cellular response to homocysteine, as in the case of a POSH complex comprising a HERPUD1. In certain embodiments, a POSH complex relates to nuclear membrane positioning and/or anchoring, as in the case of a POSH complex comprising an UNC84 polypeptide.

In an additional aspect, the application provides nucleic acid therapies for manipulating POSH. In one embodiment, the application provides a ribonucleic acid comprising between 5 and 1000 consecutive nucleotides of a nucleic acid sequence that is at least 90%, 95%, 98%, 99% or optionally 100% identical to a sequence of SEQ ID NO:1 and/or 3 or a complement thereof. Optionally the ribonucleic acid comprises at least 10, 15, 20, 25, or 30 consecutive nucleotides, and no more than 1000, 750, 500 and 250 consecutive nucleotides of a POSH nucleic

acid. In certain embodiments the ribonucleic acid is an RNAi oligomer or a ribozyme. Preferably, the ribonucleic acid decreases the level of a POSH mRNA. Preferred ribonucleic acids comprise a sequence selected from any of SEQ ID Nos: 15, 16, 18, 19, 21, 22, 24 and 25.

5 The application also features transgenic non-human animals, e.g. mice, rats, rabbits, goats, sheep, dogs, cats, cows, or non-human primates, having a transgene, e.g., animals which include (and preferably express) a heterologous form of the POSH gene described herein. Such a transgenic animal can serve as an animal model for studying viral infections such as HIV infection or for use in drug
10 screening for viral infections.

 In further aspects, the application provides compositions for the delivery of a nucleic acid therapy, such as, for example, compositions comprising a liposome and/or a pharmaceutically acceptable excipient or carrier.

 In further aspects, the application provides an isolated, purified or
15 recombinant complex comprising a POSH polypeptide and a POSH-AP. In certain
embodiments, the complex comprises a POSH-AP that interacts with a POSH
polypeptide in a yeast two-hybrid assay. In certain aspects a POSH-AP is selected
from the group consisting of: UNC48B, MSTP028 and HERPUD1 (or portion
thereof sufficient for POSH interaction). In certain embodiments, the complex
20 comprises a POSH-AP that co-immunoprecipitates with a POSH polypeptide. In
certain aspects a POSH-AP is selected from the group consisting of: dynamin II,
GroEL, an HSP70, an HSC70, a cytokeratin I, a keratin type II subunit, a cytokeratin
10, a ubiquitin-specific protease (hydrolase) and a Gag polypeptide (such as a a
Gag-Pol polypeptide). In certain embodiments, a POSH-AP is selected from the
25 group consisting of MLK1, MLK2, MLK3, MKK4, MKK7, JNK1 and JNK2. In
certain embodiments, the POSH polypeptide is a POSH RING domain, such as the
RING domain of SEQ ID NO:26 or a polypeptide at least 90% identical to SEQ ID
NO:26. In certain preferred embodiment, the complex comprises a POSH RING
domain and a polypeptide selected from the group consisting of: an HSP70-8, a Gag
30 polypeptide (such as a Gag-Pol polypeptide) and ubiquitin-specific protease
(hydrolase). In certain embodiments, the POSH polypeptide is a POSH SH3
domain, such as the SH3₄ domain of SEQ ID NO:30 or a polypeptide at least 90%

identical to SEQ ID NO:30. In certain embodiments, a complex comprises a POSH polypeptide lacking a RING domain and a polypeptide selected from the group consisting of: UNC48B, MSTP028 and HERPUD1. In certain aspects, the application provides methods for identifying a test agent having antiviral or anti-
 5 apoptotic activities by identifying a test agent that disrupts a complex described above.

In certain aspects, the application provides for identifying an agent with antiviral activity, the method comprising: a) forming a mixture comprising a test agent, a POSH polypeptide and a Gag polypeptide under conditions that, but for the
 10 test agent, permit POSH-mediated ubiquitination of the Gag polypeptide; and b) detecting POSH-mediated ubiquitination of the Gag polypeptide, wherein a decrease in POSH-mediated ubiquitination of the Gag polypeptide is indicative of an agent having antiviral or antiapoptotic activity. In preferred embodiments, the Gag polypeptide is a Gag-Pol polypeptide, and particularly an HIV p160 Gag-Pol
 15 polypeptide.

In certain aspects, the application provides methods for identifying an antiviral or antiapoptotic agent comprising: a) providing a POSH-AP polypeptide and a test agent; and b) identifying a test agent that binds to the POSH-AP polypeptide. In certain aspects the method comprises a) contacting a POSH-AP
 20 polypeptide with a test agent, and b) identifying a test agent that modulates an activity of the POSH-AP. Preferred POSH-APs for use in such a method include UNC84B, MSTP028 and HERPUD1.

In certain aspects the application provides an isolated antibody, or fragment thereof, specifically immunoreactive with an epitope of a sequence selected from the
 25 group consisting of SEQ ID NO: 2, SEQ ID No: 26, and SEQ ID NO:30, which antibody disrupts the interaction between a polypeptide of SEQ ID NO: 2 and a POSH-AP. In a preferred embodiment, the antibody or fragment thereof disrupts the interaction between a POSH domain and a POSH-AP selected from the group consisting of: UNC84B, MSTP028 and HERPUD1.

30 In certain aspects, the application provides methods of inhibiting viral infections comprising administering an agent to a subject in need thereof wherein said agent inhibits the interaction between a POSH polypeptide and a POSH-AP. In

certain preferred embodiments, the POSH-AP is a Gag polypeptide, particularly a Gag-Pol polypeptide such as the HIV p160 Gag-Pol polypeptide. Optionally, the infection is an HIV infection. In certain aspects the application provides methods of inhibiting viral infection comprising administering an agent to a subject in need thereof wherein said agent inhibits the POSH-mediated ubiquitination of a Gag polypeptide, particularly a Gag-Pol polypeptide such as the HIV p160 Gag-Pol polypeptide.

The practice of the present application will employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, transgenic biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See, for example, *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989); *DNA Cloning*, Volumes I and II (D. N. Glover ed., 1985); *Oligonucleotide Synthesis* (M. J. Gait ed., 1984); Mullis et al. U.S. Patent No: 4,683,195; *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds. 1984); *Transcription And Translation* (B. D. Hames & S. J. Higgins eds. 1984); *Culture Of Animal Cells* (R. I. Freshney, Alan R. Liss, Inc., 1987); *Immobilized Cells And Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide To Molecular Cloning* (1984); the treatise, *Methods In Enzymology* (Academic Press, Inc., N.Y.); *Gene Transfer Vectors For Mammalian Cells* (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor Laboratory); *Methods In Enzymology*, Vols. 154 and 155 (Wu et al. eds.), *Immunochemical Methods In Cell And Molecular Biology* (Mayer and Walker, eds., Academic Press, London, 1987); *Handbook Of Experimental Immunology*, Volumes I-IV (D. M. Weir and C. C. Blackwell, eds., 1986); *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986).

Other features and advantages of the application will be apparent from the following detailed description, and from the claims.

30 BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1: Human POSH Coding Sequence (SEQ ID NO:1)

Figure 2: Human POSH Amino Acid Sequence (SEQ ID NO:2)

Figure 3: Human POSH cDNA Sequence (SEQ ID NO:3)

Figure 4: 5' cDNA fragment of human POSH (public gi:10432611; SEQ ID NO:4)

Figure 5: N terminus protein fragment of hPOSH (public gi:10432612; SEQ ID NO:5)

5 Figure 6: 3' mRNA fragment of hPOSH (public gi:7959248; SEQ ID NO:6)

Figure 7: C terminus protein fragment of hPOSH (public gi:7959249; SEQ ID NO:7)

Figure 8: Human POSH full mRNA, annotated sequence

Figure 9: Domain analysis of human POSH

10 Figure 10: Diagram of human POSH nucleic acids. The diagram shows the full-length POSH gene and the position of regions amplified by RT-PCR or targeted by siRNA used in figure 11.

Figure 11: Knockdown of POSH mRNA by siRNA duplexes. HeLa SS-6 cells were transfected with siRNA against Lamin A/C (lanes 1, 2) or POSH (lanes 3-10).

15 POSH siRNA was directed against the coding region (153 - lanes 3,4; 155 - lanes 5,6) or the 3'UTR (157 - lanes 7, 8; 159 - lanes 9, 10). Cells were harvested 24 hours post-transfection, RNA extracted, and POSH mRNA levels compared by RT-PCR of a discrete sequence in the coding region of the POSH gene (see figure 10). GAPDH is used as an RT-PCR control in each reaction.

20 Figure 12: POSH affects the release of VLP from cells. A) Phosphorimages of SDS-PAGE gels of immunoprecipitations of 35S pulse-chase labeled Gag proteins are presented for cell and viral lysates from transfected HeLa cells that were either untreated or treated with POSH RNAi (50nM for 48 hours). The time during the chase period (1,2,3,4 and 5 hours after the pulse) are presented from left to right for
25 each image.

Figure 13: Release of VLP from cells at steady state. HeLa cells were transfected with an HIV-encoding plasmid and siRNA. Lanes 1, 3 and 4 were transfected with wild-type HIV-encoding plasmid. Lane 2 was transfected with an HIV-encoding plasmids which contains a point mutation in p6 (PTAP to ATAP). Control siRNA (lamin A/C) was transfected to cells in lanes 1 and 2. siRNA to Tsg101 was
30 transfected in lane 4 and siRNA to POSH in lane 3.

Figure 14: Mouse POSH mRNA sequence (public gi:10946921; SEQ ID NO: 8)

Figure 15: Mouse POSH Protein sequence (Public gi: 10946922; SEQ ID NO: 9)

Figure 16: *Drosophila melanogaster* POSH mRNA sequence (public gi:17737480; SEQ ID NO:10)

Figure 17: *Drosophila melanogaster* POSH protein sequence (public gi:17737481; SEQ ID NO:11)

Figure 18: POSH Domain Analysis

Figure 19: Partial knockdown of human POSH results in four logs reduction of HIV1 infectivity. The results from infectivity assay are presented in the diagram. The vertical axis shows the percentage of target cells infected, and the horizontal axis shows the fold dilution of virus stocks used (see Example 4 for details of the experiment). The open squares (top line) indicate the results from the control, and the closed squares (bottom line) indicate the results from transfecting cells with RNAi to POSH.

Figure 20: Human POSH protein mediates both PTAP- and PPEY-mediated viral particle release in the context of HIV. Phosphorimages of SDS-PAGE gels of immunoprecipitations of 35S pulse-chase labeled Gag proteins are presented for cell and viral lysates from transfected HeLa cells. HeLa cells were transfected with an HIV-encoding plasmid and siRNA. From left to right: the left panel was transfected with wild-type HIV-encoding plasmid. The middle panel was transfected with an HIV-encoding plasmids which contains a point mutation (PPEE to PPEY). The right panel was transfected with an HIV-encoding plasmids which contains a point mutation in p6 (PTAP to ATAP). From top to bottom: control siRNA (lamin A/C) was transfected to cells in the top panel. siRNA to Tsg101 was transfected to cells in the middle panel and siRNA to POSH was transfected to cells in the bottom panel. The time during the chase period (0, 1, 2, 4 and 8 hours after the pulse) are presented from left to right for each panel.

Figure 21: Deletion of the RING domain and the 4th SH3 domain of human POSH protein inhibits HIV1 viral particle release.

Figure 22: Human POSH has ubiquitin ligase activity

Figure 23: Human POSH co-immunoprecipitates with RAC1

Figure 24: Knock-down of human POSH entraps HIV virus particles in intracellular vesicles. HIV virus release was analyzed by electron microscopy following siRNA

and full-length HIV plasmid transfection. Mature viruses were secreted by cells transfected with HIV plasmid and non-relevant siRNA (control, bottom panel). Knockdown of Tsg101 protein resulted in a budding defect, the viruses that were released had an immature phenotype (top panel). Knockdown of hPOSH levels
 5 resulted in accumulation of viruses inside the cell in intracellular vesicles (middle panel).

Figure 25: Human POSH is localized at two sites: one at the nucleus and one at the golgi (C)) . After HIV transfection POSH golgi localization is enhanced(D). HeLa and 293T cells were transfected with pNLenv-1. 24 hours post transfection the cells
 10 were incubated with primary antibodies against POSH and either with anti-p24 Gag (B) or an organelle marker as follows: BiP- endoplasmatic reticulum (data not shown), GM130- golgi, (C,D) nucleoporin- nuclei matrix (E, F) and histone H1- nucleus (data not shown). Two different fluorescence labeled secondary antibodies were used. Specific signals were obtained by laser-scanning confocal microscopy,
 15 and individual signals were overlaid to assess their relative positions. POSH is localized to the nucleus (A) and partially colocalized with HIV-1 Gag and the golgi marker outside the nucleus, presumably in the nuclear matrix (B) following transfection.

Figures 26 and 27 show a schematic protocol and results, respectively, for
 20 identifying proteins that are ubiquitinated through a POSH-mediated process.

DETAILED DESCRIPTION OF THE APPLICATION

1. Definitions

The term "binding" refers to a direct association between two molecules, due
 25 to, for example, covalent, electrostatic, hydrophobic, ionic and/or hydrogen-bond interactions under physiological conditions.

A "chimeric protein" or "fusion protein" is a fusion of a first amino acid sequence encoding a polypeptide with a second amino acid sequence defining a domain foreign to and not substantially homologous with any domain of the first
 30 amino acid sequence. A chimeric protein may present a foreign domain which is found (albeit in a different protein) in an organism which also expresses the first

protein, or it may be an "interspecies", "intergenic", etc. fusion of protein structures expressed by different kinds of organisms.

The terms "compound", "test compound" and "molecule" are used herein interchangeably and are meant to include, but are not limited to, peptides, nucleic
 5 acids, carbohydrates, small organic molecules, natural product extract libraries, and any other molecules (including, but not limited to, chemicals, metals and organometallic compounds).

The phrase "conservative amino acid substitution" refers to grouping of amino acids on the basis of certain common properties. A functional way to define
 10 common properties between individual amino acids is to analyze the normalized frequencies of amino acid changes between corresponding proteins of homologous organisms (Schulz, G. E. and R. H. Schirmer., Principles of Protein Structure, Springer-Verlag). According to such analyses, groups of amino acids may be defined where amino acids within a group exchange preferentially with each other,
 15 and therefore resemble each other most in their impact on the overall protein structure (Schulz, G. E. and R. H. Schirmer., Principles of Protein Structure, Springer-Verlag). Examples of amino acid groups defined in this manner include:

- (i) a charged group, consisting of Glu and Asp, Lys, Arg and His,
- (ii) a positively-charged group, consisting of Lys, Arg and His,
- 20 (iii) a negatively-charged group, consisting of Glu and Asp,
- (iv) an aromatic group, consisting of Phe, Tyr and Trp,
- (v) a nitrogen ring group, consisting of His and Trp,
- (vi) a large aliphatic nonpolar group, consisting of Val, Leu and Ile,
- (vii) a slightly-polar group, consisting of Met and Cys,
- 25 (viii) a small-residue group, consisting of Ser, Thr, Asp, Asn, Gly, Ala, Glu, Gln and Pro,
- (ix) an aliphatic group consisting of Val, Leu, Ile, Met and Cys, and
- (x) a small hydroxyl group consisting of Ser and Thr.

In addition to the groups presented above, each amino acid residue may form
 30 its own group, and the group formed by an individual amino acid may be referred to simply by the one and/or three letter abbreviation for that amino acid commonly used in the art.

A "conserved residue" is an amino acid that is relatively invariant across a range of similar proteins. Often conserved residues will vary only by being replaced with a similar amino acid, as described above for "conservative amino acid substitution".

5 The term "domain" as used herein refers to a region of a protein that comprises a particular structure and/or performs a particular function.

The term "envelope virus" as used herein refers to any virus that uses cellular membrane and/or any organelle membrane in the viral release process.

"Homology" or "identity" or "similarity" refers to sequence similarity
10 between two peptides or between two nucleic acid molecules. Homology and identity can each be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When an equivalent position in the compared sequences is occupied by the same base or amino acid, then the molecules are identical at that position; when the equivalent site occupied by the same or a
15 similar amino acid residue (e.g., similar in steric and/or electronic nature), then the molecules can be referred to as homologous (similar) at that position. Expression as a percentage of homology/similarity or identity refers to a function of the number of identical or similar amino acids at positions shared by the compared sequences. A sequence which is "unrelated" or "non-homologous" shares less than 40% identity,
20 though preferably less than 25% identity with a sequence of the present application. In comparing two sequences, the absence of residues (amino acids or nucleic acids) or presence of extra residues also decreases the identity and homology/similarity.

The term "homology" describes a mathematically based comparison of sequence similarities which is used to identify genes or proteins with similar
25 functions or motifs. The nucleic acid and protein sequences of the present application may be used as a "query sequence" to perform a search against public databases to, for example, identify other family members, related sequences or homologs. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, et al. (1990) J Mol. Biol. 215:403-10. BLAST
30 nucleotide searches can be performed with the NBLAST program, score=100, wordlength=12 to obtain nucleotide sequences homologous to nucleic acid molecules of the application. BLAST protein searches can be performed with the

XBLAST program, score=50, wordlength=3 to obtain amino acid sequences homologous to protein molecules of the application. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al., (1997) *Nucleic Acids Res.* 25(17):3389-3402. When utilizing BLAST and

5 Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and BLAST) can be used. See <http://www.ncbi.nlm.nih.gov>.

As used herein, "identity" means the percentage of identical nucleotide or amino acid residues at corresponding positions in two or more sequences when the sequences are aligned to maximize sequence matching, i.e., taking into account gaps

10 and insertions. Identity can be readily calculated by known methods, including but not limited to those described in (Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A. M., and Griffin, H. G., eds., Humana

15 Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., *SIAM J. Applied Math.*, 48: 1073 (1988). Methods to determine identity are designed to give the largest match between the sequences tested. Moreover, methods to

20 determine identity are codified in publicly available computer programs. Computer program methods to determine identity between two sequences include, but are not limited to, the GCG program package (Devereux, J., et al., *Nucleic Acids Research* 12(1): 387 (1984)), BLASTP, BLASTN, and FASTA (Altschul, S. F. et al., *J. Molec. Biol.* 215: 403-410 (1990) and Altschul et al. *Nuc. Acids Res.* 25: 3389-3402

25 (1997)). The BLAST X program is publicly available from NCBI and other sources (BLAST Manual, Altschul, S., et al., NCBI NLM NIH Bethesda, Md. 20894; Altschul, S., et al., *J. Mol. Biol.* 215: 403-410 (1990). The well known Smith Waterman algorithm may also be used to determine identity.

The term "intron" refers to a portion of nucleic acid that is initially

30 transcribed into RNA but later removed such that it is not, for the most part, represented in the processed mRNA. Intron removal occurs through reactions at the 5' and 3' ends, typically referred to as 5' and 3' splice sites, respectively. Alternate

use of different splice sites results in splice variants. An intron is not necessarily situated between two "exons", or portions that code for amino acids, but may instead be positioned, for example, between the promoter and the first exon. An intron may be self-splicing or may require cellular components to be spliced out of the mRNA.

5 A "heterologous intron" is an intron that is inserted into a coding sequence that is not naturally associated with that coding sequence. In addition, a heterologous intron may be a generally natural intron wherein one or both of the splice sites have been altered to provide a desired quality, such as increased or decreased splice efficiency. Heterologous introns are often inserted, for example, to improve
10 expression of a gene in a heterologous host, or to increase the production of one splice variant relative to another. As an example, the rabbit beta-globin gene may be used, and is commercially available on the pCI vector from Promega Inc. Other exemplary introns are provided in Lacy-Hulbert et al. (2001) Gene Ther 8(8):649-53.

15 The term "isolated", as used herein with reference to the subject proteins and protein complexes, refers to a preparation of protein or protein complex that is essentially free from contaminating proteins that normally would be present with the protein or complex, e.g., in the cellular milieu in which the protein or complex is found endogenously. Thus, an isolated protein complex is isolated from cellular
20 components that normally would "contaminate" or interfere with the study of the complex in isolation, for instance while screening for modulators thereof. It is to be understood, however, that such an "isolated" complex may incorporate other proteins the modulation of which, by the subject protein or protein complex, is being investigated.

25 The term "isolated" as also used herein with respect to nucleic acids, such as DNA or RNA, refers to molecules in a form which does not occur in nature. Moreover, an "isolated nucleic acid" is meant to include nucleic acid fragments which are not naturally occurring as fragments and would not be found in the natural state.

30 Lentiviruses include primate lentiviruses, e.g., human immunodeficiency virus types 1 and 2 (HIV-1/HIV-2); simian immunodeficiency virus (SIV) from Chimpanzee (SIVcpz), Sooty mangabey (SIVsmm), African Green Monkey

(SIVagm), Syke's monkey (SIVsyk), Mandrill (SIVmnd) and Macaque (SIVmac). Lentiviruses also include feline lentiviruses, e.g., Feline immunodeficiency virus (FIV); Bovine lentiviruses, e.g., Bovine immunodeficiency virus (BIV); Ovine lentiviruses, e.g., Maedi/Visna virus (MVV) and Caprine arthritis encephalitis virus (CAEV); and Equine lentiviruses, e.g., Equine infectious anemia virus (EIAV). All lentiviruses express at least two additional regulatory proteins (Tat, Rev) in addition to Gag, Pol, and Env proteins. Primate lentiviruses produce other accessory proteins including Nef, Vpr, Vpu, Vpx, and Vif. Generally, lentiviruses are the causative agents of a variety of disease, including, in addition to immunodeficiency, neurological degeneration, and arthritis. Nucleotide sequences of the various lentiviruses can be found in Genbank under the following Accession Nos. (from J. M. Coffin, S. H. Hughes, and H. E. Varmus, "Retroviruses" Cold Spring Harbor Laboratory Press, 199,7 p 804): 1) HIV-1: K03455, M19921, K02013, M3843 1, M38429, K02007 and M17449; 2) HIV-2: M30502, J04542, M30895, J04498, M15390, M31113 and L07625; 3) SIV:M29975, M30931, M58410, M66437, L06042, M33262, M19499, M32741, M31345 and L03295; 4) FIV: M25381, M36968 and U1 1820; 5)BIV. M32690; 6)E1AV: M16575, M87581 and U01866; 6)Visna: M10608, M51543, L06906, M60609 and M60610; 7) CAEV: M33677; and 8) Ovine lentivirus M31646 and M34193. Lentiviral DNA can also be obtained from the American Type Culture Collection (ATCC). For example, feline immunodeficiency virus is available under ATCC Designation No. VR-2333 and VR-3112. Equine infectious anemia virus A is available under ATCC Designation No. VR-778. Caprine arthritis-encephalitis virus is available under ATCC Designation No. VR-905. Visna virus is available under ATCC Designation No. VR-779. As used herein, the term "nucleic acid" refers to polynucleotides such as deoxyribonucleic acid (DNA), and, where appropriate, ribonucleic acid (RNA). The term should also be understood to include, as equivalents, analogs of either RNA or DNA made from nucleotide analogs, and, as applicable to the embodiment being described, single-stranded (such as sense or antisense) and double-stranded polynucleotides.

The term "maturation" as used herein refers to the production, post-translational processing, assembly and/or release of proteins that form a viral

particle. Accordingly, this includes the processing of viral proteins leading to the pinching off of nascent virion from the cell membrane.

5 A "membrane associated protein" is meant to include proteins that are integral membrane proteins as well as proteins that are stably associated with a membrane.

The term "p6" or p6gag" is used herein to refer to a protein comprising a viral L domain. Antibodies that bind to a p6 domain are referred to as "anti-p6 antibodies". p6 also refers to proteins that comprise artificially engineered L domains including, for example, L domains comprising a series of L motifs.

10 The term "Gag protein" or "Gag polypeptide" refers to a polypeptide having Gag activity and preferably comprising an L (or late) domain. Exemplary Gag proteins include a motif such as PXXP, PPXY, RXXPXXP, RPDPTAP, RPLPVAP, RPEPTAP, YEDL, PTAPPEY and/or RPEPTAPPEE. HIV p24 is an exemplary Gag polypeptide. Gag-Pol proteins, such as the HIV p160 Gag-Pol are also Gag
15 proteins

A "POSH nucleic acid" is a nucleic acid comprising a sequence as represented in any of SEQ ID Nos: 1, 3, 4, 6, 8, and 10 as well as any of the variants described herein.

20 A "POSH polypeptide" or "POSH protein" is a polypeptide comprising a sequence as represented in any of SEQ ID Nos: 2, 5, 7, 9 and 11 as well as any of the variations described herein.

A "POSH-associated protein" or "POSH-AP" refers to a protein capable of interacting with and/or binding to a POSH polypeptide. Generally, the POSH-AP may interact directly or indirectly with the POSH polypeptide. Exemplary POSH-
25 APs are provided throughout.

The terms peptides, proteins and polypeptides are used interchangeably herein.

30 The term "purified protein" refers to a preparation of a protein or proteins which are preferably isolated from, or otherwise substantially free of, other proteins normally associated with the protein(s) in a cell or cell lysate. The term "substantially free of other cellular proteins" (also referred to herein as "substantially free of other contaminating proteins") is defined as encompassing individual

preparations of each of the component proteins comprising less than 20% (by dry weight) contaminating protein, and preferably comprises less than 5% contaminating protein. Functional forms of each of the component proteins can be prepared as purified preparations by using a cloned gene as described in the attached examples.

5 By "purified", it is meant, when referring to component protein preparations used to generate a reconstituted protein mixture, that the indicated molecule is present in the substantial absence of other biological macromolecules, such as other proteins (particularly other proteins which may substantially mask, diminish, confuse or alter the characteristics of the component proteins either as purified preparations or in

10 their function in the subject reconstituted mixture). The term "purified" as used herein preferably means at least 80% by dry weight, more preferably in the range of 85% by weight, more preferably 95-99% by weight, and most preferably at least 99.8% by weight, of biological macromolecules of the same type present (but water, buffers, and other small molecules, especially molecules having a molecular weight

15 of less than 5000, can be present). The term "pure" as used herein preferably has the same numerical limits as "purified" immediately above.

A "receptor" or "protein having a receptor function" is a protein that interacts with an extracellular ligand or a ligand that is within the cell but in a space that is topologically equivalent to the extracellular space (eg. inside the Golgi, inside

20 the endoplasmic reticulum, inside the nuclear membrane, inside a lysosome or transport vesicle, etc.). Exemplary receptors are identified herein by annotation as such in various public databases. Receptors often have membrane domains.

A "recombinant nucleic acid" is any nucleic acid that has been placed adjacent to another nucleic acid by recombinant DNA techniques. A "recombined

25 nucleic acid" also includes any nucleic acid that has been placed next to a second nucleic acid by a laboratory genetic technique such as, for example, transformation and integration, transposon hopping or viral insertion. In general, a recombined nucleic acid is not naturally located adjacent to the second nucleic acid.

The term "recombinant protein" refers to a protein of the present application

30 which is produced by recombinant DNA techniques, wherein generally DNA encoding the expressed protein is inserted into a suitable expression vector which is in turn used to transform a host cell to produce the heterologous protein. Moreover,

the phrase "derived from", with respect to a recombinant gene encoding the recombinant protein is meant to include within the meaning of "recombinant protein" those proteins having an amino acid sequence of a native protein, or an amino acid sequence similar thereto which is generated by mutations including substitutions and deletions of a naturally occurring protein.

A "RING domain" or "Ring Finger" is a zinc-binding domain with a defined octet of cysteine and histidine residues. Certain RING domains comprise the consensus sequences as set forth below (amino acid nomenclature is as set forth in Table 1): Cys Xaa Xaa Cys Xaa₁₀₋₂₀ Cys Xaa His Xaa₂₋₅ Cys Xaa Xaa Cys Xaa₁₃₋₅₀ Cys Xaa Xaa Cys or Cys Xaa Xaa Cys Xaa₁₀₋₂₀ Cys Xaa His Xaa₂₋₅ His Xaa Xaa Cys Xaa₁₃₋₅₀ Cys Xaa Xaa Cys. Certain RING domains are represented as amino acid sequences that are at least 80% identical to amino acids 12-52 of SEQ ID NO: 2 and is set forth in SEQ ID No: 26. Preferred RING domains are 85%, 90%, 95%, 98% and, most preferably, 100% identical to the amino acid sequence of SEQ ID NO: 26. Preferred RING domains of the application bind to various protein partners to form a complex that has ubiquitin ligase activity. RING domains preferably interact with at least one of the following protein types: F box proteins, E2 ubiquitin conjugating enzymes and cullins.

The term "RNA interference" or "RNAi" refers to any method by which expression of a gene or gene product is decreased by introducing into a target cell one or more double-stranded RNAs which are homologous to the gene of interest (particularly to the messenger RNA of the gene of interest).

"Small molecule" as used herein, is meant to refer to a composition, which has a molecular weight of less than about 5 kD and most preferably less than about 2.5 kD. Small molecules can be nucleic acids, peptides, polypeptides, peptidomimetics, carbohydrates, lipids or other organic (carbon containing) or inorganic molecules. Many pharmaceutical companies have extensive libraries of chemical and/or biological mixtures comprising arrays of small molecules, often fungal, bacterial, or algal extracts, which can be screened with any of the assays of the application.

An "SH3" or "Src Homology 3" domain is a protein domain of generally about 60 amino acid residues first identified as a conserved sequence in the non-

catalytic part of several cytoplasmic protein tyrosine kinases (e.g. Src, Abl, Lck). SH3 domains mediate assembly of specific protein complexes via binding to proline-rich peptides. Exemplary SH3 domains are represented by amino acids 137-192, 199-258, 448-505 and 832-888 of SEQ ID NO:2 and are set forth in SEQ ID
 5 Nos: 27-30. In certain embodiments, an SH3 domain interacts with a consensus sequence of RXaaXaaPXaaX6P (where X6, as defined in table 1 below, is a hydrophobic amino acid). In certain embodiments, an SH3 domain interacts with one or more of the following sequences: P(T/S)AP, PFRDY, RPEPTAP, RQGPKEP, RQGPKEPFR, RPEPTAPEE and RPLPVAP.

10 As used herein, the term "specifically hybridizes" refers to the ability of a nucleic acid probe/primer of the application to hybridize to at least 12, 15, 20, 25, 30, 35, 40, 45, 50 or 100 consecutive nucleotides of a POSH sequence, or a sequence complementary thereto, or naturally occurring mutants thereof, such that it has less than 15%, preferably less than 10%, and more preferably less than 5%
 15 background hybridization to a cellular nucleic acid (e.g., mRNA or genomic DNA) other than the POSH gene. A variety of hybridization conditions may be used to detect specific hybridization, and the stringency is determined primarily by the wash stage of the hybridization assay. Generally high temperatures and low salt concentrations give high stringency, while low temperatures and high salt
 20 concentrations give low stringency. Low stringency hybridization is achieved by washing in, for example, about 2.0 x SSC at 50 °C, and high stringency is achieved with about 0.2 x SSC at 50 °C. Further descriptions of stringency are provided below.

As applied to polypeptides, "substantial sequence identity" means that two
 25 peptide sequences, when optimally aligned, such as by the programs GAP or BESTFIT using default gap which share at least 90 percent sequence identity, preferably at least 95 percent sequence identity, more preferably at least 99 percent sequence identity or more. Preferably, residue positions which are not identical differ by conservative amino acid substitutions. For example, the substitution of
 30 amino acids having similar chemical properties such as charge or polarity are not likely to effect the properties of a protein. Examples include glutamine for asparagine or glutamic acid for aspartic acid.

"Transcriptional regulatory sequence" is a generic term used throughout the specification to refer to DNA sequences, such as initiation signals, enhancers, and promoters, which induce or control transcription of protein coding sequences with which they are operably linked. In preferred embodiments, transcription of a recombinant protein gene is under the control of a promoter sequence (or other transcriptional regulatory sequence) which controls the expression of the recombinant gene in a cell-type in which expression is intended. It will also be understood that the recombinant gene can be under the control of transcriptional regulatory sequences which are the same or which are different from those sequences which control transcription of the naturally-occurring form of the protein.

As used herein, a "transgenic animal" is any animal, preferably a non-human mammal, bird or an amphibian, in which one or more of the cells of the animal contain heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. This molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA. In the typical transgenic animals described herein, the transgene causes cells to express a recombinant human POSH protein. The "non-human animals" of the application include vertebrates such as rodents, non-human primates, sheep, dog, cow, chickens, amphibians, reptiles, etc. Preferred non-human animals are selected from the rodent family including rat and mouse, most preferably mouse, though transgenic amphibians, such as members of the *Xenopus* genus, and transgenic chickens can also provide important tools for understanding and identifying agents which can affect, for example, embryogenesis and tissue formation. The term "chimeric animal" is used herein to refer to animals in which the recombinant gene is found, or in which the recombinant is expressed in some but not all cells of the animal. The term "tissue specific chimeric animal" indicates that the recombinant human POSH genes is present and/or expressed in some tissues but not others.

As used herein, the term "transgene" means a nucleic acid sequence (encoding, e.g., human POSH polypeptides), which is partly or entirely heterologous, i.e., foreign, to the transgenic animal or cell into which it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but which is designed to be inserted, or is inserted, into the animal's genome in such a way as to alter the genome of the cell into which it is inserted (e.g., it is inserted at a location which differs from that of the natural gene or its insertion results in a knockout). A transgene can include one or more transcriptional regulatory sequences and any other nucleic acid, such as introns, that may be necessary for optimal expression of a selected nucleic acid.

As is well known, genes for a particular polypeptide may exist in single or multiple copies within the genome of an individual. Such duplicate genes may be identical or may have certain modifications, including nucleotide substitutions, additions or deletions, which all still code for polypeptides having substantially the same activity.

A "virion" is a complete viral particle; nucleic acid and capsid (and a lipid envelope in some viruses).

*Table 1: Abbreviations for classes of amino acids**

Symbol	Category	Amino Acids Represented
X1	Alcohol	Ser, Thr
X2	Aliphatic	Ile, Leu, Val
Xaa	Any	Ala, Cys, Asp, Glu, Phe, Gly, His, Ile, Lys, Leu, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp, Tyr

X4	Aromatic	Phe, His, Trp, Tyr
X5	Charged	Asp, Glu, His, Lys, Arg
X6	Hydrophobic	Ala, Cys, Phe, Gly, His, Ile, Lys, Leu, Met, Thr, Val, Trp, Tyr
X7	Negative	Asp, Glu
X8	Polar	Cys, Asp, Glu, His, Lys, Asn, Gln, Arg, Ser, Thr
X9	Positive	His, Lys, Arg
X10	Small	Ala, Cys, Asp, Gly, Asn, Pro, Ser, Thr, Val
X11	Tiny	Ala, Gly, Ser
X12	Turnlike	Ala, Cys, Asp, Glu, Gly, His, Lys, Asn, Gln, Arg, Ser, Thr
X13	Asparagine-Aspartate	Asn, Asp

* Abbreviations as adopted from http://smart.embl-heidelberg.de/SMART_DATA/alignments/consensus/grouping.html.

2. Overview

- 5 In certain aspects, the application relates to the discovery of novel associations between POSH proteins and other proteins (termed POSH-APs), and related methods and compositions. In certain aspects, the application relates to novel associations between certain disease states and POSH nucleic acids and

proteins. In certain aspects, the application relates to novel associations between certain disease states and POSH-AP nucleic acids and proteins.

POSH intersects with and regulates a wide range of key cellular functions that may be manipulated by affecting the level of and/or activity of POSH polypeptides or POSH-AP polypeptides. In certain aspects, by identifying proteins associated with POSH, and particularly human POSH the present application provides methods for identifying diseases that are associated with defects in the POSH gene and methods for ameliorating such diseases. In further aspects, the application provides nucleic acid agents (e.g. RNAi probes, antisense), antibody-related agents, small molecules and other agents that affect POSH function. In further aspects, the application provides methods for identifying agents that affect POSH function, and the function of proteins that associate with POSH and/or participate in a POSH or POSH-AP mediated process. Other aspects and embodiments are described herein.

In certain aspects, the application relates to the discovery that certain POSH polypeptides function as E3 enzymes in the ubiquitination system. Accordingly, downregulation or upregulation of POSH ubiquitin ligase activity can be used to manipulate biological processes that are affected by protein ubiquitination. Downregulation or upregulation may be achieved at any stage of POSH formation and regulation, including transcriptional, translational or post-translational regulation. For example, POSH transcript levels may be decreased by RNAi targeted at a POSH gene sequence. As another example, POSH ubiquitin ligase activity may be inhibited by contacting POSH with an antibody that binds to and interferes with a POSH RING domain or a domain of POSH that mediates interaction with a target protein (a protein that is ubiquitinated at least in part because of POSH activity). As another example, POSH activity may be increased by causing increased expression of POSH or an active portion thereof. A ubiquitin ligase, such as POSH, and POSH-APs that modulate the POSH ubiquitin ligase activity may participate in biological processes including, for example, one or more of the various stages of a viral lifecycle, such as viral entry into a cell, production of viral proteins, assembly of viral proteins and release of viral particles from the cell. POSH may participate in diseases characterized by the accumulation of

ubiquitinated proteins, such as dementias (e.g. Alzheimer's and Pick's), inclusion body myositis and myopathies, polyglucosan body myopathy, and certain forms of amyotrophic lateral sclerosis. POSH may participate in diseases characterized by the excessive or inappropriate ubiquitination and/or protein degradation. In addition, POSH may participate in oncological processes, such as the failure of cell division control systems, the failure of cell death regulatory systems, and the failure to downregulate hyperactive oncogenes, such as hyperactive membrane-bound growth factor receptors. By identifying certain POSH polypeptides as ubiquitin ligases, aspects of the present application permit one of ordinary skill in the art to identify diseases that are associated with an altered POSH ubiquitin ligase activity.

In certain aspects, the application relates to the discovery that certain POSH polypeptides and POSH-APs are involved in viral maturation, including the production, post-translational processing, assembly and/or release of proteins in a viral particle. Accordingly, viral infections may be ameliorated by inhibiting an activity (e.g. ubiquitin ligase activity or target protein interaction) of POSH, and in preferred embodiments, the virus is a retroid virus, an RNA virus and an envelope virus, including HIV, Ebola, HBV, HCV and HTLV. Additional viral species are described in greater detail below. In certain instances, a decrease of a POSH function is lethal to cells infected with a virus that employs POSH in release of viral particles. While not wishing to be bound to mechanism, it appears that loss of POSH function in such cells leads to cell death through an overaccumulation of viral particles, or portions thereof, in a the cell. In certain embodiments, the inhibition of a POSH activity, e.g. by siRNA knockdown of POSH or certain POSH-APs may be used to destroy infected cells, even cells with nearly latent virus, because such cells will die from eventual overaccumulation of viral particles or portions thereof.

In certain aspects, the application relates to the discovery that hPOSH interacts with Rac, a small GTPase. Rho, Rac and Cdc42 operate together to regulate organization of the actin cytoskeleton and the MLK-MKK-JNK MAP kinase pathway (Xu et al., 2003, EMBO J. 2: 252-61). Ectopic expression of mouse POSH ("mPOSH") activates the JNK pathway and causes nuclear localization of NF- κ B. Overexpression of mPOSH in fibroblasts stimulates apoptosis. (Tapon et al. (1998) EMBO J. 17:1395-404). In Drosophila, POSH may interact, or otherwise

influence the signaling of, another GTPase, Ras. (Schnorr et al. (2001) Genetics 159: 609-22). The JNK pathway and NF- κ B regulate a variety of key genes involved in, for example, immune responses, inflammation, cell proliferation and apoptosis. For example, NF- κ B regulates the production of interleukin 1, interleukin 8, tumor necrosis factor and many cell adhesion molecules. NF- κ B has both pro-apoptotic and anti-apoptotic roles in the cell (e.g. in FAS-induced cell death and TNF-alpha signaling, respectively). NF- κ B is negatively regulated, in part, by the inhibitor proteins I κ B α and I κ B β (collectively termed "I κ B"). Phosphorylation of I κ B permits activation and nuclear localization of NF- κ B.

10 Phosphorylation of I κ B triggers its degradation by the ubiquitin system. Accordingly, in yet another embodiment, a POSH polypeptide stimulates the JNK pathway. In an additional embodiment, a POSH polypeptide promotes nuclear localization of NF- κ B. In further embodiments, manipulation of POSH levels and/or activities may be used to manipulate apoptosis. By upregulating POSH,

15 apoptosis may be stimulated in certain cells, and this will generally be desirable in conditions characterized by excessive cell proliferation (e.g. in certain cancers). By downregulating POSH, apoptosis may be diminished in certain cells, and this will generally be desirable in conditions characterized by excessive cell death, such as myocardial infarction, stroke, degenerative diseases of muscle and nerve, and for

20 organ preservation prior to transplant. In a further embodiment, a POSH polypeptide associates with a vesicular trafficking complex, such as a clathrin- or coatamer- containing complex, and particularly a trafficking complex that localizes to the nucleus and/or Golgi apparatus.

In certain aspects, the application relates to the discovery that a POSH

25 polypeptide interacts with human UNC84B, a human homolog of *C. elegans* Unc-84. In *C. elegans*, Unc-84 is involved in the cellular positioning of the nucleus. UNC84/SUN is positioned at the nuclear membrane and recruits Syne/ANC-1, which directly tethers the nuclear envelope to the actin cytoskeleton. Accordingly, in certain aspects, POSH participates in formation of a UNC84 complexes, including

30 human UNC84B-containing complexes, and in the connections between the nucleus and the cytoskeleton. In certain aspects, UNC84 polypeptides participate in POSH-

mediated processes. See, for example, Starr and Han, 2003, J Cell Sci 116(Pt 2):211-6. The term UNC84 is used herein to refer to various naturally occurring Unc-84 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring UNC84. The term specifically includes human UNC84B nucleic acid and amino acid sequences and the sequences presented in the Examples.

In certain aspects, the application relates to the discovery that a POSH polypeptide interacts with MSTP028. Certain MSTP028 polypeptides contain one or more BTB/POZ domains that are generally involved in dimerization. Accordingly the application provides complexes comprising POSH and MSTP028, optionally in a dimeric form. The term MSTP028 is used herein to refer to various naturally occurring MSTP028 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring MSTP028. The term specifically includes human MSTP028 nucleic acid and amino acid sequences and the sequences presented in the Examples.

In certain aspects, the application relates to the discovery that a POSH polypeptide interacts with HERPUD1, a "homocysteine-inducible, endoplasmic reticulum stress-inducible, ubiquitin-like domain member 1" protein. (In previous applications, HERPUD1 was referred to as HEPUD1). Certain HERPUD1 polypeptides are involved in JNK-mediated apoptosis, particularly in vascular endothelial cells, including cells that are exposed to high levels of homocysteine. Certain HERPUD1 polypeptides are involved in the Unfolded Protein Response, a cellular response to the presence of unfolded proteins in the endoplasmic reticulum. Certain HERPUD1 polypeptides are involved in the regulation of sterol biosynthesis. Accordingly, certain POSH polypeptides are involved in the Unfolded Protein Response and sterol biosynthesis. In other aspects, certain HERPUD1 polypeptides enhance presenilin-mediated amyloid β -protein generation. For example, HERPUD1 polypeptides, when overexpressed in cells, increase the level of amyloid β generation, and it is observed that HERPUD1 polypeptides interact with the presenilin proteins, presenilin-1 and presenilin-2. (See Sai, X. et al (2002) J. Biol. Chem. 277:12915-12920). Accordingly, in certain aspects, POSH

polypeptides may modulate the level of amyloid β generation. Additionally, POSH polypeptides may interact with presenilin 1 and presenilin 2. Therefore, it is believed certain POSH polypeptides modulate presenilin-mediated amyloid β generation. The accumulation of amyloid β is one hallmark of Alzheimer's disease.

5 Accordingly, these POSH polypeptides may be involved in the pathogenesis of Alzheimer's disease. At sites such as late intracellular compartment sites including the trans-Golgi network, certain mutant presenilin-2 polypeptides up-regulate production of amyloid β peptides ending at position 42 (A β 42). (See Iwata, H. et al (2001) J. Biol. Chem. 276: 21678-21685). Accordingly, POSH polypeptides

10 regulate production of A β 42 through mutant presenilin-2 at late intracellular compartment sites including the trans-Golgi network. Furthermore, elevated homocysteine levels have been found to be a risk factor associated with Alzheimer's disease and cerebral vascular disease. Some risk factors, such as elevated plasma homocysteine levels, may accelerate or increase the severity of several central

15 nervous system (CNS) disorders. Elevated levels of plasma homocysteine were found in young male patients with schizophrenia suggesting that elevated homocysteine levels could be related to the pathophysiology of aspects of schizophrenia (Levine, J. et al (2002) Am. J. Psychiatry 159:1790-2). Accordingly, certain POSH polypeptides may be involved in neurological disorders. Neurological

20 disorders include disorders associated with increased levels of plasma homocysteine, increased levels of amyloid β production, or aberrant presenilin activity. Neurological disorders include CNS disorders, such as Alzheimer's disease, cerebral vascular disease and schizophrenia. Certain POSH polypeptides may be involved in cardiovascular diseases, such as thromboembolic vascular disease, and particularly

25 the disease characteristics associated with hyperhomocysteinemia. See, for example, Kokame et al. 2000 J. Biol. Chem. 275:32846-53; Zhang et al. 2001 Biochem Biophys Res Commun 289:718-24. The term HERPUD1 is used herein to refer to various naturally occurring HERPUD1 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence

30 identity to a naturally occurring HERPUD1. The term specifically includes human HERPUD1 nucleic acid and amino acid sequences and the sequences presented in the Examples.

3. Exemplary Nucleic Acids and Expression Vectors

In certain aspects the application provides nucleic acids encoding POSH polypeptides, such as, for example, SEQ ID Nos: 2, 5, 7, 9, 11, 26, 27, 28, 29 and 30. Nucleic acids of the application are further understood to include nucleic acids that comprise variants of SEQ ID Nos: 1, 3, 4, 6, 8, 10, 31, 32, 33, 34, and 35. Variant nucleotide sequences include sequences that differ by one or more nucleotide substitutions, additions or deletions, such as allelic variants; and will, therefore, include coding sequences that differ from the nucleotide sequence of the coding sequence designated in SEQ ID Nos: 1, 3, 4, 6, 8, 10, 31, 32, 33, 34, and 35, e.g., due to the degeneracy of the genetic code. In other embodiments, variants will also include sequences that will hybridize under highly stringent conditions to a nucleotide sequence of a coding sequence designated in any of SEQ ID Nos: 1, 3, 4, 6, 8, 10, 31, 32, 33, 34, and 35. Preferred nucleic acids of the application are human POSH sequences, including, for example, any of SEQ ID Nos: 1, 3, 4, 6, 31, 32, 33, 34, 35 and variants thereof and nucleic acids encoding an amino acid sequence selected from among SEQ ID Nos: 2, 5, 7, 26, 27, 28, 29 and 30.

One of ordinary skill in the art will understand readily that appropriate stringency conditions which promote DNA hybridization can be varied. For example, one could perform the hybridization at 6.0 x sodium chloride/sodium citrate (SSC) at about 45 °C, followed by a wash of 2.0 x SSC at 50 °C. For example, the salt concentration in the wash step can be selected from a low stringency of about 2.0 x SSC at 50 °C to a high stringency of about 0.2 x SSC at 50 °C. In addition, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22 °C, to high stringency conditions at about 65 °C. Both temperature and salt may be varied, or temperature or salt concentration may be held constant while the other variable is changed. In one embodiment, the application provides nucleic acids which hybridize under low stringency conditions of 6 x SSC at room temperature followed by a wash at 2 x SSC at room temperature.

Isolated nucleic acids which differ from SEQ ID Nos: 1, 3, 4, 6, 8, 10, 31, 32, 33, 34, and 35 due to degeneracy in the genetic code are also within the scope of the application. For example, a number of amino acids are designated by more than one

triplet. Codons that specify the same amino acid, or synonyms (for example, CAU and CAC are synonyms for histidine) may result in "silent" mutations which do not affect the amino acid sequence of the protein. However, it is expected that DNA sequence polymorphisms that do lead to changes in the amino acid sequences of the subject proteins will exist among mammalian cells. One skilled in the art will appreciate that these variations in one or more nucleotides (up to about 3-5% of the nucleotides) of the nucleic acids encoding a particular protein may exist among individuals of a given species due to natural allelic variation. Any and all such nucleotide variations and resulting amino acid polymorphisms are within the scope of this application.

Optionally, a POSH nucleic acid of the application will genetically complement a partial or complete POSH loss of function phenotype in a cell. For example, a POSH nucleic acid of the application may be expressed in a cell in which endogenous POSH has been reduced by RNAi, and the introduced POSH nucleic acid will mitigate a phenotype resulting from the RNAi. An exemplary POSH loss of function phenotype is a decrease in virus-like particle production in a cell transfected with a viral vector, optionally an HIV vector. In certain embodiments, a POSH nucleic acid, when expressed at an effective level in a cell, induces apoptosis.

Another aspect of the application relates to POSH nucleic acids that are used for antisense, RNAi or ribozymes. As used herein, nucleic acid therapy refers to administration or *in situ* generation of a nucleic acid or a derivative thereof which specifically hybridizes (e.g. binds) under cellular conditions with the cellular mRNA and/or genomic DNA encoding one of the subject POSH polypeptides so as to inhibit production of that protein, e.g. by inhibiting transcription and/or translation. The binding may be by conventional base pair complementarity, or, for example, in the case of binding to DNA duplexes, through specific interactions in the major groove of the double helix.

An nucleic acid therapy construct of the present application can be delivered, for example, as an expression plasmid which, when transcribed in the cell, produces RNA which is complementary to at least a unique portion of the cellular mRNA which encodes a POSH polypeptide. Alternatively, the the construct is an oligonucleotide which is generated *ex vivo* and which, when introduced into the cell

causes inhibition of expression by hybridizing with the mRNA and/or genomic sequences encoding a POSH polypeptide. Such oligonucleotide probes are optionally modified oligonucleotide which are resistant to endogenous nucleases, e.g. exonucleases and/or endonucleases, and is therefore stable *in vivo*. Exemplary

5 nucleic acid molecules for use as antisense oligonucleotides are phosphoramidate, phosphothioate and methylphosphonate analogs of DNA (see also U.S. Patents 5,176,996; 5,264,564; and 5,256,775). Additionally, general approaches to constructing oligomers useful in nucleic acid therapy have been reviewed, for example, by van der Krol et al., (1988) *Biotechniques* 6:958-976; and Stein et al.,

10 (1988) *Cancer Res* 48:2659-2668.

Accordingly, the modified oligomers of the application are useful in therapeutic, diagnostic, and research contexts. In therapeutic applications, the oligomers are utilized in a manner appropriate for nucleic acid therapy in general.

In addition to use in therapy, the oligomers of the application may be used as

15 diagnostic reagents to detect the presence or absence of the POSH DNA or RNA sequences to which they specifically bind, such as for determining the level of expression of a gene of the application or for determining whether a gene of the application contains a genetic lesion.

In another aspect of the application, the subject nucleic acid is provided in an

20 expression vector comprising a nucleotide sequence encoding a subject POSH polypeptide and operably linked to at least one regulatory sequence. Regulatory sequences are art-recognized and are selected to direct expression of the POSH polypeptide. Accordingly, the term regulatory sequence includes promoters, enhancers and other expression control elements. Exemplary regulatory sequences

25 are described in Goeddel; *Gene Expression Technology: Methods in Enzymology*, Academic Press, San Diego, CA (1990). For instance, any of a wide variety of expression control sequences that control the expression of a DNA sequence when operatively linked to it may be used in these vectors to express DNA sequences encoding a POSH polypeptide. Such useful expression control sequences, include,

30 for example, the early and late promoters of SV40, tet promoter, adenovirus or cytomegalovirus immediate early promoter, the lac system, the trp system, the TAC or TRC system, T7 promoter whose expression is directed by T7 RNA polymerase,

the major operator and promoter regions of phage lambda , the control regions for fd coat protein, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase, e.g., Pho5, the promoters of the yeast α -mating factors, the polyhedron promoter of the baculovirus system and other
5 sequences known to control the expression of genes of prokaryotic or eukaryotic cells or their viruses, and various combinations thereof. It should be understood that the design of the expression vector may depend on such factors as the choice of the host cell to be transformed and/or the type of protein desired to be expressed. Moreover, the vector's copy number, the ability to control that copy number and the
10 expression of any other protein encoded by the vector, such as antibiotic markers, should also be considered.

As will be apparent, the subject gene constructs can be used to cause expression of the subject POSH polypeptides in cells propagated in culture, e.g. to produce proteins or polypeptides, including fusion proteins or polypeptides, for
15 purification.

This application also pertains to a host cell transfected with a recombinant gene including a coding sequence for one or more of the subject POSH polypeptides. The host cell may be any prokaryotic or eukaryotic cell. For example, a polypeptide of the present application may be expressed in bacterial cells
20 such as *E. coli*, insect cells (e.g., using a baculovirus expression system), yeast, or mammalian cells. Other suitable host cells are known to those skilled in the art.

Accordingly, the present application further pertains to methods of producing the subject POSH polypeptides. For example, a host cell transfected with an expression vector encoding a POSH polypeptide can be cultured under
25 appropriate conditions to allow expression of the polypeptide to occur. The polypeptide may be secreted and isolated from a mixture of cells and medium containing the polypeptide. Alternatively, the polypeptide may be retained cytoplasmically and the cells harvested, lysed and the protein isolated. A cell culture includes host cells, media and other byproducts. Suitable media for cell
30 culture are well known in the art. The polypeptide can be isolated from cell culture medium, host cells, or both using techniques known in the art for purifying proteins, including ion-exchange chromatography, gel filtration chromatography,

ultrafiltration, electrophoresis, and immunoaffinity purification with antibodies specific for particular epitopes of the polypeptide. In a preferred embodiment, the POSH polypeptide is a fusion protein containing a domain which facilitates its purification, such as a POSH-GST fusion protein, POSH-intein fusion protein,
 5 POSH-cellulose binding domain fusion protein, POSH-polyhistidine fusion protein etc.

A nucleotide sequence encoding a POSH polypeptide can be used to produce a recombinant form of the protein via microbial or eukaryotic cellular processes. Ligating the polynucleotide sequence into a gene construct, such as an expression
 10 vector, and transforming or transfecting into hosts, either eukaryotic (yeast, avian, insect or mammalian) or prokaryotic (bacterial) cells, are standard procedures.

A recombinant POSH nucleic acid can be produced by ligating the cloned gene, or a portion thereof, into a vector suitable for expression in either prokaryotic cells, eukaryotic cells, or both. Expression vehicles for production of a recombinant
 15 POSH polypeptides include plasmids and other vectors. For instance, suitable vectors for the expression of a POSH polypeptide include plasmids of the types: pBR322-derived plasmids, pEMBL-derived plasmids, pEX-derived plasmids, pBTac-derived plasmids and pUC-derived plasmids for expression in prokaryotic cells, such as *E. coli*.

20 A number of vectors exist for the expression of recombinant proteins in yeast. For instance, YEP24, YIP5, YEP51, YEP52, pYES2, and YRP17 are cloning and expression vehicles useful in the introduction of genetic constructs into *S. cerevisiae* (see, for example, Broach et al., (1983) in *Experimental Manipulation of Gene Expression*, ed. M. Inouye Academic Press, p.
 25 83, incorporated by reference herein). These vectors can replicate in *E. coli* due the presence of the pBR322 ori, and in *S. cerevisiae* due to the replication determinant of the yeast 2 micron plasmid. In addition, drug resistance markers such as ampicillin can be used.

The preferred mammalian expression vectors contain both prokaryotic
 30 sequences to facilitate the propagation of the vector in bacteria, and one or more eukaryotic transcription units that are expressed in eukaryotic cells. The pcDNA1/amp, pcDNA1/neo, pRc/CMV, pSV2gpt, pSV2neo, pSV2-dhfr, pTk2,

pRSVneo, pMSG, pSVT7, pko-neo and pHyg derived vectors are examples of mammalian expression vectors suitable for transfection of eukaryotic cells. Some of these vectors are modified with sequences from bacterial plasmids, such as pBR322, to facilitate replication and drug resistance selection in both prokaryotic and eukaryotic cells. Alternatively, derivatives of viruses such as the bovine papilloma virus (BPV-1), or Epstein-Barr virus (pHEBo, pREP-derived and p205) can be used for transient expression of proteins in eukaryotic cells. Examples of other viral (including retroviral) expression systems can be found below in the description of gene therapy delivery systems. The various methods employed in the preparation of the plasmids and transformation of host organisms are well known in the art. For other suitable expression systems for both prokaryotic and eukaryotic cells, as well as general recombinant procedures, see *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press, 1989) Chapters 16 and 17. In some instances, it may be desirable to express the recombinant POSH polypeptide by the use of a baculovirus expression system. Examples of such baculovirus expression systems include pVL-derived vectors (such as pVL1392, pVL1393 and pVL941), pAcUW-derived vectors (such as pAcUW1), and pBlueBac-derived vectors (such as the β -gal containing pBlueBac III).

It is well known in the art that a methionine at the N-terminal position can be enzymatically cleaved by the use of the enzyme methionine aminopeptidase (MAP). MAP has been cloned from *E. coli* (Ben-Bassat et al., (1987) *J. Bacteriol.* 169:751-757) and *Salmonella typhimurium* and its *in vitro* activity has been demonstrated on recombinant proteins (Miller et al., (1987) *PNAS USA* 84:2718-1722). Therefore, removal of an N-terminal methionine, if desired, can be achieved either *in vivo* by expressing such recombinant polypeptides in a host which produces MAP (e.g., *E. coli* or CM89 or *S. cerevisiae*), or *in vitro* by use of purified MAP (e.g., procedure of Miller et al.).

Alternatively, the coding sequences for the polypeptide can be incorporated as a part of a fusion gene including a nucleotide sequence encoding a different polypeptide. This type of expression system can be useful under conditions where it is desirable, e.g., to produce an immunogenic fragment of a POSH polypeptide. For

example, the VP6 capsid protein of rotavirus can be used as an immunologic carrier protein for portions of polypeptide, either in the monomeric form or in the form of a viral particle. The nucleic acid sequences corresponding to the portion of the POSH polypeptide to which antibodies are to be raised can be incorporated into a fusion gene construct which includes coding sequences for a late vaccinia virus structural protein to produce a set of recombinant viruses expressing fusion proteins comprising a portion of the protein as part of the virion. The Hepatitis B surface antigen can also be utilized in this role as well. Similarly, chimeric constructs coding for fusion proteins containing a portion of a POSH polypeptide and the poliovirus capsid protein can be created to enhance immunogenicity (see, for example, EP Publication NO: 0259149; and Evans et al., (1989) *Nature* 339:385; Huang et al., (1988) *J. Virol.* 62:3855; and Schlienger et al., (1992) *J. Virol.* 66:2).

The Multiple Antigen Peptide system for peptide-based immunization can be utilized, wherein a desired portion of a POSH polypeptide is obtained directly from organo-chemical synthesis of the peptide onto an oligomeric branching lysine core (see, for example, Posnett et al., (1988) *JBC* 263:1719; and Nardelli et al., (1992) *J. Immunol.* 148:914). Antigenic determinants of a POSH polypeptide can also be expressed and presented by bacterial cells.

In another embodiment, a fusion gene coding for a purification leader sequence, such as a poly-(His)/enterokinase cleavage site sequence at the N-terminus of the desired portion of the recombinant protein, can allow purification of the expressed fusion protein by affinity chromatography using a Ni^{2+} metal resin. The purification leader sequence can then be subsequently removed by treatment with enterokinase to provide the purified POSH polypeptide (e.g., see Hochuli et al., (1987) *J. Chromatography* 411:177; and Janknecht et al., *PNAS USA* 88:8972).

Techniques for making fusion genes are well known. Essentially, the joining of various DNA fragments coding for different polypeptide sequences is performed in accordance with conventional techniques, employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated

- DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel et al., John Wiley & Sons: 1992).
- 5

Table 2: Exemplary POSH nucleic acids

<u>Sequence Name</u>	<u>Organism</u>	<u>Accession Number</u>
cDNA FLJ11367 fis, clone HEMBA1000303	Homo sapiens	AK021429
Plenty of SH3 domains (POSH) mRNA	Mus musculus	NM_021506
Plenty of SH3s (POSH) mRNA	Mus musculus	AF030131
Plenty of SH3s (POSH) mRNA	Drosophila melanogaster	NM_079052
Plenty of SH3s (POSH) mRNA	Drosophila melanogaster	AF220364

Table 3: Exemplary POSH polypeptides

<u>Sequence Name</u>	<u>Organism</u>	<u>Accession Number</u>
SH3 domains-containing protein POSH	Mus musculus	T09071
plenty of SH3 domains	Mus musculus	NP_067481
Plenty of SH3s; POSH	Mus musculus	AAC40070
Plenty of SH3s	Drosophila melanogaster	AAF37265

LD45365p	Drosophila melanogaster	AAK93408
POSH gene product	Drosophila melanogaster	AAF57833
Plenty of SH3s	Drosophila melanogaster	NP_523776

In addition the following Tables provide the nucleic acid sequence and related SEQ ID NOs for domains of human POSH protein and a summary of sequence identification numbers used in this application.

5 Table 4. Nucleic Acid Sequences and related SEQ ID NOs for domains in human POSH

Name of the sequence	Sequence	SEQ ID NO.
RING domain	TGTCCGGTGTGTCTAGAGCGCCTTGATGCTTCTGCGAAGGTCT TGCCTTGCCAGCATACGTTTTGCAAGCGATGTTTGCT GGGGATCGTAGGTTCTCGAATGAACTCAGATGTCCCGAGT	31
1 st SH ₃ domain	CCATGTGCCAAAGCGTTATACAATATGAAGGAAAAGAGCCTG GAGACCTTAAATTCAGCAAAGGCGACATCATCATTTT GCGAAGACAAGTGGATGAAAATTGGTACCATGGGGAAGTCAAT GGAATCCATGGCTTTTTCCCCACCAACTTTGTGCAGA TTATT	32
2 nd SH ₃ domain	CCTCAGTGCAAAGCACTTTATGACTTTGAAGTGAAAGACAAGG AAGCAGACAAAGATTGCCTTCCATTGCAAAGGATGA TGTTCTGACTGTGATCCGAAGAGTGGATGAAAAGTGGGCTGAA GGAATGCTGGCAGACAAAATAGGAATATTTCCAATTT CATATGTTGAGTTTAAC	33
3 rd SH ₃ domain	AGTGTGTATGTTGCTATATATCCATACACTCCTCGGAAAGAGG ATGAACTAGAGCTGAGAAAAGGGGAGATGTTTTTAGT GTTTGAGCGCTGCCAGGATGGCTGGTTCAAAGGGACATCCATG CATACCAGCAAGATAGGGGTTTTCCCTGGCAATTATG	34

	TGGCACCAGTC	
4 th SH ₃ domain	GAAAGGCACAGGGTGGTGGTTTCCTATCCTCCTCAGAGTGAGG CAGAACTTGAACCTAAAGAAGGAGATATTGTGTTTGT TCATAAAAAACGAGAGGATGGCTGGTTCAAAGGCACATTACAA CGTAATGGGAAAAC TGGCCTTTTCCCAGGAAGCTTTG TGGAAAACA	35

Table 5. Summary of Sequence Identification Numbers

Sequence Information	Sequence Identification Number (SEQ ID NO)
Human POSH Coding Sequence	SEQ ID No: 1
Human POSH Amino Acid Sequence	SEQ ID No: 2
Human POSH cDNA Sequence	SEQ ID No: 3
5' cDNA Fragment of Human POSH	SEQ ID No: 4
N-terminus Protein Fragment of Human POSH	SEQ ID No: 5
3' mRNA Fragment of Human POSH	SEQ ID No: 6
C-terminus Protein Fragment of Human POSH	SEQ ID No: 7
Mouse POSH mRNA Sequence	SEQ ID No: 8
Mouse POSH Protein Sequence	SEQ ID No: 9
Drosophila melanogaster POSH mRNA Sequence	SEQ ID No: 10
Drosophila melanogaster POSH Protein Sequence	SEQ ID No: 11
Human POSH RING Domain Amino Acid Sequence	SEQ ID No: 26
Human POSH 1 st SH ₃ Domain Amino Acid Sequence	SEQ ID No: 27
Human POSH 2 nd SH ₃ Domain Amino Acid Sequence	SEQ ID No: 28
Human POSH 3 rd SH ₃ Domain Amino Acid Sequence	SEQ ID No: 29
Human POSH 4 th SH ₃ Domain Amino	SEQ ID No: 30

Acid Sequence	
Human POSH RING Domain Nucleic Acid Sequence	SEQ ID No: 31
Human POSH 1 st SH ₃ Domain Nucleic Acid Sequence	SEQ ID No: 32
Human POSH 2 nd SH ₃ Domain Nucleic Acid Sequence	SEQ ID No: 33
Human POSH 3 rd SH ₃ Domain Nucleic Acid Sequence	SEQ ID No: 34
Human POSH 4 th SH ₃ Domain Nucleic Acid Sequence	SEQ ID No: 35

4. Exemplary Polypeptides

The present application also makes available isolated and/or purified forms of the subject POSH polypeptides, which are isolated from, or otherwise substantially free of, other intracellular proteins which might normally be associated with the protein or a particular complex including the protein. In certain embodiments, POSH polypeptides have an amino acid sequence that is at least 60% identical to an amino acid sequence as set forth in any of SEQ ID Nos: 2, 5, 7, 9, 11, 26, 27, 28, 29 and 30. In other embodiments, the polypeptide has an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99% or 100% identical to an amino acid sequence as set forth in any of SEQ ID Nos: 2, 5, 7, 9, 11, 26, 27, 28, 29 and 30.

Optionally, a POSH polypeptide of the application will function in place of an endogenous POSH polypeptide, for example by mitigating a partial or complete POSH loss of function phenotype in a cell. For example, a POSH polypeptide of the application may be produced in a cell in which endogenous POSH has been reduced by RNAi, and the introduced POSH polypeptide will mitigate a phenotype resulting from the RNAi. An exemplary POSH loss of function phenotype is a decrease in virus-like particle production in a cell transfected with a viral vector, optionally an HIV vector. In certain embodiments, a POSH polypeptide, when produced at an effective level in a cell, induces apoptosis.

In certain embodiments, a POSH polypeptide of the application interacts with a viral Gag protein. In additional embodiments, POSH polypeptides may also,

or alternatively, function in ubiquitylation in part through the activity of a RING domain. In certain embodiments, POSH interacts with Gag polypeptides, and particularly Gag-Pol polypeptides, through ubiquitination mediated by the POSH RING domain.

5 In another aspect, the application provides polypeptides that are agonists or antagonists of a POSH polypeptide. Variants and fragments of a POSH polypeptide may have a hyperactive or constitutive activity, or, alternatively, act to prevent POSH polypeptides from performing one or more functions. For example, a truncated form lacking one or more domain may have a dominant negative effect.

10 Another aspect of the application relates to polypeptides derived from a full-length POSH polypeptide. Isolated peptidyl portions of the subject proteins can be obtained by screening polypeptides recombinantly produced from the corresponding fragment of the nucleic acid encoding such polypeptides. In addition, fragments can be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry. For example, any one of the
15 subject proteins can be arbitrarily divided into fragments of desired length with no overlap of the fragments, or preferably divided into overlapping fragments of a desired length. The fragments can be produced (recombinantly or by chemical synthesis) and tested to identify those peptidyl fragments which can function as
20 either agonists or antagonists of the formation of a specific protein complex, or more generally of a POSH complex, such as by microinjection assays.

It is also possible to modify the structure of the subject POSH polypeptides for such purposes as enhancing therapeutic or prophylactic efficacy, or stability (e.g., ex vivo shelf life and resistance to proteolytic degradation in vivo). Such
25 modified polypeptides, when designed to retain at least one activity of the naturally-occurring form of the protein, are considered functional equivalents of the POSH polypeptides described in more detail herein. Such modified polypeptides can be produced, for instance, by amino acid substitution, deletion, or addition.

For instance, it is reasonable to expect, for example, that an isolated
30 replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, or a similar replacement of an amino acid with a structurally related amino acid (i.e. conservative mutations) will not have a major

effect on the biological activity of the resulting molecule. Conservative replacements are those that take place within a family of amino acids that are related in their side chains. Genetically encoded amino acids can be divided into four families: (1) acidic = aspartate, glutamate; (2) basic = lysine, arginine, histidine; (3) nonpolar = alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan; and (4) uncharged polar = glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine. Phenylalanine, tryptophan, and tyrosine are sometimes classified jointly as aromatic amino acids. In similar fashion, the amino acid repertoire can be grouped as (1) acidic = aspartate, glutamate; (2) basic = lysine, arginine, histidine, (3) aliphatic = glycine, alanine, valine, leucine, isoleucine, serine, threonine, with serine and threonine optionally be grouped separately as aliphatic-hydroxyl; (4) aromatic = phenylalanine, tyrosine, tryptophan; (5) amide = asparagine, glutamine; and (6) sulfur-containing = cysteine and methionine. (see, for example, Biochemistry, 2nd ed., Ed. by L. Stryer, W.H. Freeman and Co., 1981). Whether a change in the amino acid sequence of a polypeptide results in a functional homolog can be readily determined by assessing the ability of the variant polypeptide to produce a response in cells in a fashion similar to the wild-type protein. For instance, such variant forms of a POSH polypeptide can be assessed, e.g., for their ability to bind to another polypeptide, e.g., another POSH polypeptide or another protein involved in viral maturation. Polypeptides in which more than one replacement has taken place can readily be tested in the same manner.

This application further contemplates a method of generating sets of combinatorial mutants of the subject POSH polypeptides, as well as truncation mutants, and is especially useful for identifying potential variant sequences (e.g. homologs) that are functional in binding to a POSH polypeptide. The purpose of screening such combinatorial libraries is to generate, for example, POSH homologs which can act as either agonists or antagonist, or alternatively, which possess novel activities all together. Combinatorially-derived homologs can be generated which have a selective potency relative to a naturally occurring POSH polypeptide. Such proteins, when expressed from recombinant DNA constructs, can be used in gene therapy protocols.

Likewise, mutagenesis can give rise to homologs which have intracellular half-lives dramatically different than the corresponding wild-type protein. For example, the altered protein can be rendered either more stable or less stable to proteolytic degradation or other cellular process which result in destruction of, or otherwise inactivation of the POSH polypeptide of interest. Such homologs, and the genes which encode them, can be utilized to alter POSH levels by modulating the half-life of the protein. For instance, a short half-life can give rise to more transient biological effects and, when part of an inducible expression system, can allow tighter control of recombinant POSH levels within the cell. As above, such proteins, and particularly their recombinant nucleic acid constructs, can be used in gene therapy protocols.

In similar fashion, POSH homologs can be generated by the present combinatorial approach to act as antagonists, in that they are able to interfere with the ability of the corresponding wild-type protein to function.

In a representative embodiment of this method, the amino acid sequences for a population of POSH homologs are aligned, preferably to promote the highest homology possible. Such a population of variants can include, for example, homologs from one or more species, or homologs from the same species but which differ due to mutation. Amino acids which appear at each position of the aligned sequences are selected to create a degenerate set of combinatorial sequences. In a preferred embodiment, the combinatorial library is produced by way of a degenerate library of genes encoding a library of polypeptides which each include at least a portion of potential POSH sequences. For instance, a mixture of synthetic oligonucleotides can be enzymatically ligated into gene sequences such that the degenerate set of potential POSH nucleotide sequences are expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g. for phage display).

There are many ways by which the library of potential homologs can be generated from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be carried out in an automatic DNA synthesizer, and the synthetic genes then be ligated into an appropriate gene for expression. The purpose of a degenerate set of genes is to provide, in one mixture, all of the

sequences encoding the desired set of potential POSH sequences. The synthesis of degenerate oligonucleotides is well known in the art (see for example, Narang, SA (1983) *Tetrahedron* 39:3; Itakura et al., (1981) *Recombinant DNA*, Proc. 3rd Cleveland Sympos. Macromolecules, ed. AG Walton, Amsterdam: Elsevier pp273-289; Itakura et al., (1984) *Annu. Rev. Biochem.* 53:323; Itakura et al., (1984) *Science* 198:1056; Ike et al., (1983) *Nucleic Acid Res.* 11:477). Such techniques have been employed in the directed evolution of other proteins (see, for example, Scott et al., (1990) *Science* 249:386-390; Roberts et al., (1992) *PNAS USA* 89:2429-2433; Devlin et al., (1990) *Science* 249: 404-406; Cwirla et al., (1990) *PNAS USA* 87: 6378-6382; as well as U.S. Patent Nos: 5,223,409, 5,198,346, and 5,096,815).

Alternatively, other forms of mutagenesis can be utilized to generate a combinatorial library. For example, POSH homologs (both agonist and antagonist forms) can be generated and isolated from a library by screening using, for example, alanine scanning mutagenesis and the like (Ruf et al., (1994) *Biochemistry* 33:1565-1572; Wang et al., (1994) *J. Biol. Chem.* 269:3095-3099; Balint et al., (1993) *Gene* 137:109-118; Grodberg et al., (1993) *Eur. J. Biochem.* 218:597-601; Nagashima et al., (1993) *J. Biol. Chem.* 268:2888-2892; Lowman et al., (1991) *Biochemistry* 30:10832-10838; and Cunningham et al., (1989) *Science* 244:1081-1085), by linker scanning mutagenesis (Gustin et al., (1993) *Virology* 193:653-660; Brown et al., (1992) *Mol. Cell Biol.* 12:2644-2652; McKnight et al., (1982) *Science* 232:316); by saturation mutagenesis (Meyers et al., (1986) *Science* 232:613); by PCR mutagenesis (Leung et al., (1989) *Method Cell Mol Biol* 1:11-19); or by random mutagenesis, including chemical mutagenesis, etc. (Miller et al., (1992) *A Short Course in Bacterial Genetics*, CSHL Press, Cold Spring Harbor, NY; and Greener et al., (1994) *Strategies in Mol Biol* 7:32-34). Linker scanning mutagenesis, particularly in a combinatorial setting, is an attractive method for identifying truncated (bioactive) forms of POSH polypeptides.

A wide range of techniques are known in the art for screening gene products of combinatorial libraries made by point mutations and truncations, and, for that matter, for screening cDNA libraries for gene products having a certain property. Such techniques will be generally adaptable for rapid screening of the gene libraries

generated by the combinatorial mutagenesis of POSH homologs. The most widely used techniques for screening large gene libraries typically comprises cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates relatively easy isolation of the vector encoding the gene whose product was detected. Each of the illustrative assays described below are amenable to high through-put analysis as necessary to screen large numbers of degenerate sequences created by combinatorial mutagenesis techniques.

10 In an illustrative embodiment of a screening assay, candidate combinatorial gene products of one of the subject proteins are displayed on the surface of a cell or virus, and the ability of particular cells or viral particles to bind a POSH polypeptide is detected in a "panning assay". For instance, a library of POSH variants can be cloned into the gene for a surface membrane protein of a bacterial cell (Ladner et al., WO 88/06630; Fuchs et al., (1991) *Bio/Technology* 9:1370-1371; and Goward et al., (1992) *TIBS* 18:136-140), and the resulting fusion protein detected by panning, e.g. using a fluorescently labeled molecule which binds the POSH polypeptide, to score for potentially functional homologs. Cells can be visually inspected and separated under a fluorescence microscope, or, where the morphology of the cell permits, separated by a fluorescence-activated cell sorter.

20 In similar fashion, the gene library can be expressed as a fusion protein on the surface of a viral particle. For instance, in the filamentous phage system, foreign peptide sequences can be expressed on the surface of infectious phage, thereby conferring two significant benefits. First, since these phage can be applied to affinity matrices at very high concentrations, a large number of phage can be screened at one time. Second, since each infectious phage displays the combinatorial gene product on its surface, if a particular phage is recovered from an affinity matrix in low yield, the phage can be amplified by another round of infection. The group of almost identical *E. coli* filamentous phages M13, fd, and f1 are most often used in phage display libraries, as either of the phage gIII or gVIII coat proteins can be used to generate fusion proteins without disrupting the ultimate packaging of the viral particle (Ladner et al., PCT publication WO 90/02909;

Garrard et al., PCT publication WO 92/09690; Marks et al., (1992) J. Biol. Chem. 267:16007-16010; Griffiths et al., (1993) EMBO J. 12:725-734; Clackson et al., (1991) Nature 352:624-628; and Barbas et al., (1992) PNAS USA 89:4457-4461).

The application also provides for reduction of the subject POSH polypeptides to generate mimetics, e.g. peptide or non-peptide agents, which are able to mimic binding of the authentic protein to another cellular partner. Such mutagenic techniques as described above, as well as the thioredoxin system, are also particularly useful for mapping the determinants of a POSH polypeptide which participate in protein-protein interactions involved in, for example, binding of proteins involved in viral maturation to each other. To illustrate, the critical residues of a POSH polypeptide which are involved in molecular recognition of a substrate protein can be determined and used to generate POSH polypeptide-derived peptidomimetics which bind to the substrate protein, and by inhibiting POSH binding, act to inhibit its biological activity. By employing, for example, scanning mutagenesis to map the amino acid residues of a POSH polypeptide which are involved in binding to another polypeptide, peptidomimetic compounds can be generated which mimic those residues involved in binding. For instance, non-hydrolyzable peptide analogs of such residues can be generated using benzodiazepine (e.g., see Freidinger et al., in Peptides: Chemistry and Biology, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), azepine (e.g., see Huffman et al., in Peptides: Chemistry and Biology, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), substituted gamma lactam rings (Garvey et al., in Peptides: Chemistry and Biology, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), keto-methylene pseudopeptides (Ewenson et al., (1986) J. Med. Chem. 29:295; and Ewenson et al., in Peptides: Structure and Function (Proceedings of the 9th American Peptide Symposium) Pierce Chemical Co. Rockland, IL, 1985), b-turn dipeptide cores (Nagai et al., (1985) Tetrahedron Lett 26:647; and Sato et al., (1986) J Chem Soc Perkin Trans 1:1231), and b-aminoalcohols (Gordon et al., (1985) Biochem Biophys Res Commun 126:419; and Dann et al., (1986) Biochem Biophys Res Commun 134:71).

The following table provides the sequences of the RING domain and the various SH3 domains.

Table 6. Amino Acid Sequences and related SEQ ID NOs for domains in human POSH

Name of the sequence	Sequence	SEQ ID NO.
RING domain	CPVCLERLDASAKVLPCQHTFCRCLLGIVGSRNELRCPEC	26
1 st SH ₃ domain	PCAKALYNYEGKEPGDLKFSKGDIIILRRQVDENWYHGEVNGIHGF FPTNFVQIIK	27
2 nd SH ₃ domain	PQCKALYDFEVKDKEADKDCLPFAKDDVLTVIRRV DENWAEGLAD KIGIFPISYVEFNS	28
3 rd SH ₃ domain	SVYVAIYPYTPRKEDELELRKGEMFLVFERCQDGWFKGTSMTSKI GVFPNGYVAPVT	29
4 th SH ₃ domain	ERHRVVVSYPQSEAELELKEGDIVFVHKKREDGWFKGTLQRNGKT GLFPGSFVENI	30

5

5. Antibodies and Uses Thereof

Another aspect of the application pertains to an antibody specifically reactive with a POSH polypeptide. For example, by using immunogens derived from a POSH polypeptide, e.g., based on the cDNA sequences, anti-protein/anti-peptide antisera or monoclonal antibodies can be made by standard protocols (See, for example, Antibodies: A Laboratory Manual ed. by Harlow and Lane (Cold Spring Harbor Press: 1988)). A mammal, such as a mouse, a hamster or rabbit can be immunized with an immunogenic form of the peptide (e.g., a POSH polypeptide or an antigenic fragment which is capable of eliciting an antibody response, or a fusion protein as described above). Techniques for conferring immunogenicity on a protein or peptide include conjugation to carriers or other techniques well known in the art. An immunogenic portion of a POSH polypeptide can be administered in the presence of adjuvant. The progress of immunization can be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassays can be used with the immunogen as antigen to assess the levels of antibodies. In a

preferred embodiment, the subject antibodies are immunospecific for antigenic determinants of a POSH polypeptide of a mammal, e.g., antigenic determinants of a protein set forth in SEQ ID NO:2.

In one embodiment, antibodies are specific for a RING domain or an SH3 domain, and preferably the domain is part of a POSH polypeptide. In a more specific embodiment, the domain is part of an amino acid sequence set forth in SEQ ID NO:2. In a set of exemplary embodiments, an antibody binds to one or more SH3 domains represented by amino acids 137-192 of SEQ ID NO:2, amino acids 199-258 of SEQ ID NO:2, amino acids 448-505 of SEQ ID NO:2, and/or amino acids 832-888 of SEQ ID NO:2. In another exemplary embodiment, an antibody binds to a RING domain represented by amino acids 12-52 of SEQ ID NO:2. In another embodiment, the antibodies are immunoreactive with one or more proteins having an amino acid sequence that is at least 80% identical to an amino acid sequence as set forth in SEQ ID NO:2. In other embodiments, an antibody is immunoreactive with one or more proteins having an amino acid sequence that is 85%, 90%, 95%, 98%, 99% or identical to an amino acid sequence as set forth in SEQ ID NO:2.

Following immunization of an animal with an antigenic preparation of a POSH polypeptide, anti-POSH antisera can be obtained and, if desired, polyclonal anti-POSH antibodies isolated from the serum. To produce monoclonal antibodies, antibody-producing cells (lymphocytes) can be harvested from an immunized animal and fused by standard somatic cell fusion procedures with immortalizing cells such as myeloma cells to yield hybridoma cells. Such techniques are well known in the art, and include, for example, the hybridoma technique (originally developed by Kohler and Milstein, (1975) *Nature*, 256: 495-497), the human B cell hybridoma technique (Kozbar et al., (1983) *Immunology Today*, 4: 72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., (1985) *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. pp. 77-96). Hybridoma cells can be screened immunochemically for production of antibodies specifically reactive with a mammalian POSH polypeptide of the present application and monoclonal antibodies isolated from a culture comprising such hybridoma cells.

In one embodiment anti-human POSH antibodies specifically react with the protein encoded by a nucleic acid having SEQ ID NO:2.

The term antibody as used herein is intended to include fragments thereof which are also specifically reactive with one of the subject POSH polypeptides.

- 5 Antibodies can be fragmented using conventional techniques and the fragments screened for utility in the same manner as described above for whole antibodies. For example, F(ab)₂ fragments can be generated by treating antibody with pepsin. The resulting F(ab)₂ fragment can be treated to reduce disulfide bridges to produce Fab fragments. The antibody of the present application is further intended to include
- 10 bispecific, single-chain, and chimeric and humanized molecules having affinity for a POSH polypeptide conferred by at least one CDR region of the antibody. In preferred embodiments, the antibodies, the antibody further comprises a label attached thereto and able to be detected, (e.g., the label can be a radioisotope, fluorescent compound, enzyme or enzyme co-factor).

- 15 Anti-POSH antibodies can be used, e.g., to monitor POSH polypeptide levels in an individual, particularly the presence of POSH at the plasma membrane for determining whether or not said patient is infected with a virus such as an RNA virus, a retroid virus, and an envelope virus, or allowing determination of the efficacy of a given treatment regimen for an individual afflicted with such a
- 20 disorder. In addition, POSH polypeptides are expected to localize, occasionally, to the released viral particle. Viral particles may be collected and assayed for the presence of a POSH polypeptide. The level of POSH polypeptide may be measured in a variety of sample types such as, for example, cells and/or in bodily fluid, such as in blood samples.

- 25 Another application of anti-POSH antibodies of the present application is in the immunological screening of cDNA libraries constructed in expression vectors such as gt11, gt18-23, ZAP, and ORF8. Messenger libraries of this type, having coding sequences inserted in the correct reading frame and orientation, can produce fusion proteins. For instance, gt11 will produce fusion proteins whose amino
- 30 termini consist of β -galactosidase amino acid sequences and whose carboxy termini consist of a foreign polypeptide. Antigenic epitopes of a POSH polypeptide, e.g., other orthologs of a particular protein or other paralogs from the same species, can

then be detected with antibodies, as, for example, reacting nitrocellulose filters lifted from infected plates with the appropriate anti-POSH antibodies. Positive phage detected by this assay can then be isolated from the infected plate. Thus, the presence of POSH homologs can be detected and cloned from other animals, as can
 5 alternate isoforms (including splice variants) from humans.

In certain embodiments, the application provides antibodies that disrupt the interaction between POSH and a POSH-AP. Such an antibody may bind to an epitope of the POSH-AP. Such an antibody may also bind to an epitope of POSH.

10 6. Homology Searching of Nucleotide and Polypeptide Sequences

The nucleotide or amino acid sequences of the application may be used as query sequences against databases such as GenBank, SwissProt, BLOCKS, and Pima II. These databases contain previously identified and annotated sequences that can be searched for regions of homology (similarity) using BLAST, which stands for
 15 Basic Local Alignment Search Tool (Altschul S F (1993) J Mol Evol 36:290-300; Altschul, S F et al (1990) J Mol Biol 215:403-10).

BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologs
 20 which may be of prokaryotic (bacterial) or eukaryotic (animal, fungal or plant) origin. Other algorithms such as the one described in Smith, R. F. and T. F. Smith (1992; Protein Engineering 5:35-51), incorporated herein by reference, can be used when dealing with primary sequence patterns and secondary structure gap penalties. As disclosed in this application, sequences have lengths of at least 49 nucleotides
 25 and no more than 12% uncalled bases (where N is recorded rather than A, C, G, or T).

The BLAST approach, as detailed in Karlin and Altschul (1993; Proc Nat Acad Sci 90:5873-7) and incorporated herein by reference, searches matches between a query sequence and a database sequence, to evaluate the statistical
 30 significance of any matches found, and to report only those matches which satisfy the user-selected threshold of significance. Preferably the threshold is set at 10-25 for nucleotides and 3-15 for peptides.

7. Transgenic Animals and Uses Thereof

Another aspect of the application features transgenic non-human animals which express a heterologous POSH and/or POSH-AP gene, preferentially a human POSH and/or POSH-AP gene of the present application, and/or which have had one or both copies of the endogenous POSH and/or POSH-AP genes disrupted in at least one of the tissue or cell-types of the animal. Accordingly, the application features an animal model for viral infection. In one embodiment, the transgenic non-human animals is a mammal such as a mouse, rat, rabbit, goat, sheep, dog, cat, cow, or non-human primate. Without being bound to theory, it is proposed that such an animal may be susceptible to infection with envelope viruses, retroid viruses and RNA viruses such as various rhabdoviruses, lentiviruses, and filoviruses. Accordingly, such a transgenic animal may serve as a useful animal model to study the progression of diseases caused by such viruses. Alternatively, such an animal can be useful as a basis to introduce one or more other human transgenes, to create a transgenic animal carrying multiple human genes involved in infection caused by retroid viruses, or RNA viruses, and envelope viruses. Retroid viruses include lentiviruses such as HIV. Other RNA viruses include filoviruses such as Ebola virus. As a result of the introduction of multiple human transgenes, the transgenic animal may become susceptible to certain viral infection and therefore provide an useful animal model to study these viral infection.

In a preferred embodiment, the transgenic animal carrying human POSH gene is useful as a basis to introduce other human genes involved in HIV infection, such as Cyclin T1, CD34, CCR5, and fusin (CRCX4). In a further embodiment, the additional human transgene is a gene involved in a disease or condition that is associated with AIDS (e.g. hypertension, Kaposi's sarcoma, cachexia, etc.) Such an animal may be an useful animal model for studying HIV infection, AIDS and related disease development.

Another aspect of the present application concerns transgenic animals which are comprised of cells (of that animal) which contain a transgene of the present application and which preferably (though optionally) express an exogenous POSH

and/or POSH-AP protein in one or more cells in the animal. A POSH or POSH-AP transgene can encode the wild-type form of the protein, or can encode homologs thereof, as well as antisense constructs. Moreover, it may be desirable to express the heterologous transgene conditionally such that either the timing or the level of gene
 5 expression can be regulated. Such conditional expression can be provided using prokaryotic promoter sequences which require prokaryotic proteins to be simultaneous expressed in order to facilitate expression of the transgene. Exemplary promoters and the corresponding trans-activating prokaryotic proteins are given in U.S. Pat. No. 4,833,080.

10 Moreover, transgenic animals exhibiting tissue specific expression can be generated, for example, by inserting a tissue specific regulatory element, such as an enhancer, into the transgene. For example, the endogenous POSH or POSH-AP gene promoter or a portion thereof can be replaced with another promoter and/or enhancer, e.g., a CMV or a Moloney murine leukemia virus (MLV) promoter and/or
 15 enhancer.

Alternatively, non-human transgenic animals that only express HIV transgenes in the brain can be generated using brain specific promoters (e.g. myelin basic protein (MBP) promoter, the neurofilament protein (NF-L) promoter, the gonadotropin-releasing hormone promoter, the vasopressin promoter and the
 20 neuron-specific enolase promoter, see So Forss-Petter et al., Neuron, 5, 187, (1990). Such animals can provide a useful in vivo model to evaluate the ability of a potential anti-HIV drug to cross the blood-brain barrier. Other target cells for which specific promoters can be used are, for example, macrophages, T cells and B cells. Other tissue specific promoters are well-known in the art, see e.g. R.Jaenisch, Science,
 25 240, 1468 (1988).

Non-human transgenic animals containing an inducible transgene can be generated using inducible regulatory elements (e.g. metallothionein promoter), which are well-known in the art. Transgene expression can then be initiated in these animals by administering to the animal a compound which induces gene expression
 30 (e.g. heavy metals). Another preferred inducible system comprises a tetracycline-

inducible transcriptional activator (U.S. Pat. No. 5,654,168 issued Aug. 5, 1997 to Bujard and Gossen and U.S. Pat. No. 5,650,298 issued Jul. 22, 1997 to Bujard et al.).

In general, transgenic animal lines can be obtained by generating transgenic
5 animals having incorporated into their genome at least one transgene, selecting at least one founder from these animals and breeding the founder or founders to establish at least one line of transgenic animals having the selected transgene incorporated into their genome.

Animals for obtaining eggs or other nucleated cells (e.g. embryonic stem
10 cells) for generating transgenic animals can be obtained from standard commercial sources such as Charles River Laboratories (Wilmington, Mass.), Taconic (Germantown, N.Y.), Harlan Sprague Dawley (Indianapolis, Ind.).

Eggs can be obtained from suitable animals, e.g., by flushing from the oviduct or using techniques described in U.S. Pat. No. 5,489,742 issued Feb. 6, 1996
15 to Hammer and Taurog; U.S. Pat. No. 5,625,125 issued on Apr. 29, 1997 to Bennett et al.; Gordon et al., 1980, Proc. Natl. Acad. Sci. USA 77:7380-7384; Gordon & Ruddle, 1981, Science 214: 1244-1246; U.S. Pat. No. 4,873,191 to T. E. Wagner and P. C. Hoppe; U.S. Pat. No. 5,604,131; Armstrong, et al. (1988) J. of Reproduction, 39:511 or PCT application No. PCT/FR93/00598 (WO 94/00568) by
20 Mehtali et al. Preferably, the female is subjected to hormonal conditions effective to promote superovulation prior to obtaining the eggs.

Many techniques can be used to introduce DNA into an egg or other nucleated cell, including in vitro fertilization using sperm as a carrier of exogenous DNA ("sperm-mediated gene transfer", e.g., Lavitrano et al., 1989, Cell 57: 717-
25 723), microinjection, gene targeting (Thompson et al., 1989, Cell 56: 313-321), electroporation (Lo, 1983, Mol. Cell. Biol. 3: 1803-1814), transfection, or retrovirus mediated gene transfer (Van der Putten et al., 1985, Proc. Natl. Acad. Sci. USA 82: 6148-6152). For a review of such techniques, see Gordon (1989), Transgenic Animals, Intl. Rev. Cytol. 115:171-229.

Except for sperm-mediated gene transfer, eggs should be fertilized in conjunction with (before, during or after) other transgene transfer techniques. A preferred method for fertilizing eggs is by breeding the female with a fertile male. However, eggs can also be fertilized by in vitro fertilization techniques.

5 Fertilized, transgene containing eggs can then be transferred to pseudopregnant animals, also termed "foster mother animals", using suitable techniques. Pseudopregnant animals can be obtained, for example, by placing 40-80 day old female animals, which are more than 8 weeks of age, in cages with infertile males, e.g., vasectomized males. The next morning females are checked for vaginal
10 plugs. Females who have mated with vasectomized males are held aside until the time of transfer.

Recipient females can be synchronized, e.g. using GNRH agonist (GnRH-a): des-gly10, (D-Ala6)-LH-RH Ethylamide, SigmaChemical Co., St. Louis, Mo. Alternatively, a unilateral pregnancy can be achieved by a brief surgical procedure
15 involving the "peeling" away of the bursa membrane on the left uterine horn. Injected embryos can then be transferred to the left uterine horn via the infundibulum. Potential transgenic founders can typically be identified immediately at birth from the endogenous litter mates. For generating transgenic animals from embryonic stem cells, see e.g. Teratocarcinomas and embryonic stem cells, a
20 practical approach, ed. E. J. Robertson, (IRL Press 1987) or in Potter et al Proc. Natl. Acad. Sci. USA 81, 7161 (1984), the teachings of which are incorporated herein by reference.

Founders that express the gene can then be bred to establish a transgenic line. Accordingly, founder animals can be bred, inbred, crossbred or outbred to produce
25 colonies of animals of the present application. Animals comprising multiple transgenes can be generated by crossing different founder animals (e.g. an HIV transgenic animal and a transgenic animal, which expresses human CD4), as well as by introducing multiple transgenes into an egg or embryonic cell as described above. Furthermore, embryos from A-transgenic animals can be stored as frozen embryos,
30 which are thawed and implanted into pseudo-pregnant animals when needed (See

e.g. Hirabayashi et al. (1997) Exp Anim 46: 111 and Anzai (1994) Jikken Dobutsu 43: 247).

5 The present application provides for transgenic animals that carry the transgene in all their cells, as well as animals that carry the transgene in some, but not all cells, i.e., mosaic animals. The transgene can be integrated as a single transgene or in tandem, e.g., head to head tandems, or head to tail or tail to tail or as multiple copies.

10 The successful expression of the transgene can be detected by any of several means well known to those skilled in the art. Non-limiting examples include Northern blot, in situ hybridization of mRNA analysis, Western blot analysis, immunohistochemistry, and FACS analysis of protein expression.

15 In a further aspect, the application features non-human animal cells containing a POSH or POSH-AP transgene, preferentially a human POSH or POSH-AP transgene. For example, the animal cell (e.g. somatic cell or germ cell (i.e. egg or sperm)) can be obtained from the transgenic animal. Transgenic somatic cells or cell lines can be used, for example, in drug screening assays. Transgenic germ cells, on the other hand, can be used in generating transgenic progeny, as described above.

20 The application further provides methods for identifying (screening) or for determining the safety and/or efficacy of virus therapeutics, i.e. compounds which are useful for treating and/or preventing the development of diseases or conditions, which are caused by, or contributed to by viral infection (e.g. AIDS). In addition the assays are useful for further improving known anti-viral compounds, e.g. by modifying their structure to increase their stability and/or activity and/or toxicity.

25 The transgenic animals can be used in in vivo assays to identify viral therapeutics. For example, the animals can be used in assays to identify compounds which reduce or inhibit any phase of the viral life cycle, e.g., expression of one or more viral genes, activity of one or more viral proteins, glycosylation of one or more viral proteins, processing of one or more viral proteins, viral replication, assembly of virions, and/or budding of infectious virions.

In an exemplary embodiment, the assay comprises administering a test compound to a transgenic animal of the application infected with a virus including RNA viruses, DNA viruses, retrovirus and/or envelope viruses, and comparing a phenotypic change in the animal relative to a transgenic animal which has not received the test compound. For example, where the animal is infected with HIV, the phenotypic change can be the amelioration in an AIDS related complex (ARC), cataracts, inflammatory lesions in the central nervous system (CNV), a mild kidney sclerotic lesion, or a skin lesion, such as psoriatic dermatitis, hyperkeratotic lesions, Kaposi's sarcoma or cachexia. The effect of a compound on inhibition of Kaposi's sarcoma can be determined, as described, e.g., in PCT/US97/11202 (WO97/49373) by Gallo et al. These and other HIV related symptoms or phenotypes are further described in Leonard et al. (1988) Science 242:1665.

In another embodiment, the phenotypic change is release/budding of virus particles. In yet another embodiment, the phenotypic change is the number of CD4+ T cells or the ratio of CD4+ T cells versus CD8+ T cells. In HIV infected humans as well as in HIV transgenic mice, analysis of lymph nodes indicate that the number of CD4+ T cells decreases and the number of CD8+ T cells increases. Numbers of CD4+ and CD8+ T cells can be determined, for example, by indirect immunofluorescence and flow cytometry, as described, e.g., in Santoro et al., supra.

Alternatively, a phenotypic change, e.g. a change in the expression level of an HIV gene can be monitored. The HIV RNA can be selected from the group consisting of gag mRNA, gag-pro-pol mRNA, vif mRNA, vpr mRNA, tat mRNA, rev mRNA, vpu/env mRNA, nef mRNA, and vpx mRNA. The HIV protein can be selected from the group consisting of Pr55 Gag and fragments thereof (p17 MA, p24 CA, p7 NC, p1, p9, p6, and p2), Pr160 Gag-Pro-Pol, and fragments thereof (p10 PR, p51 RT, p66 RT, p32 IN), p23 Vif, p15 Vpr, p14 Tat, p19 Rev, p16 Vpu, gPr 160 Env or fragments thereof (gp120 SU and gp41TM), p27 Nef, and p14 Vpx. The level of any of these mRNAs or proteins can be determined in cells from a tissue sample, such as a skin biopsy, as described in, e.g., PCT/US97/11202 (WO97/49373) by Gallo et al. Quantitation of HIV mRNA and protein is further described elsewhere herein and also in, e.g., Dickie et al. (1996) AIDS Res. Human

Retroviruses 12:1103. In a preferred embodiment, the level of gp120 on the surface of PBMC is determined. This can be done, as described in the examples, e.g., by immunofluorescence on PBMC obtained from the animals.

5 A further phenotypic change is the production level or rate of viral particles in the serum and/or tissue of the animal. This can be determined, e.g., by determining reverse transcriptase (RT activity) or viral load as described elsewhere herein as well as in PCT/US97/11202 (WO97/49373) by Gallo et al., such as by determining p24 antigen.

10 Yet another phenotypic change, which can indicate HIV infection or AIDS progression is the production of inflammatory cytokines such as IL-6, IL-8 and TNF- α .; thus, efficacy of a compound as an anti-HIV therapeutic can be assessed by ELISA tests for the reduction of serum levels of any or all of these cytokines.

15 A vaccine can be tested by administering a test antigen to a transgenic animal of the application. The animal can optionally be boosted with the same or a different antigen. Such animal is then infected with a virus such as HIV. The production of viral particles or expression of viral proteins is then measured at various times following the administration of the test vaccine. A decrease in the amount of viral particles produced or viral expression will indicate that the test vaccine is efficient in
20 reducing or inhibiting viral production and/or expression. The amount of antibody produced by the animal in response to the vaccine antigen can also be determined according to methods known in the art and provides a relative indication of the immunogenicity of the particular antigen.

25 Cells from the transgenic animals of the application can be established in culture and immortalized to establish cell lines. For example, immortalized cell lines can be established from the livers of transgenic rats, as described in Bulera et al. (1997) Hepatology 25: 1192. Cell lines from other types of cells can be established according to methods known in the art."

30 In one cell-based assay, cells expressing a POSH or POSH-AP transgene can be infected with a virus of interest and incubated in the presence a test compound or

a control compound. The production of viral particles is then compared. This assay system thus provides a means of identifying molecular antagonists which, for example, function by interfering with viral release/budding.

Cell based assays can also be used to identify compounds which modulate expression of a viral gene, modulate translation of a viral mRNA, or which modulate the stability of a viral mRNA or protein. Accordingly, a cell which is infected with a virus of interest can be incubated with a test compound and the amount of the viral protein produced in the cell medium can be measured and compared to that produced from a cell which has not been contacted with the test compound. The specificity of the compound for regulating the expression of the particular virus gene can be confirmed by various control analyses, e.g., measuring the expression of one or more control genes. This type of cellular assay can be particularly useful for determining the efficacy of antisense molecules or ribozymes.

8. RNA Interference, Ribozymes, Antisense and DNA Enzyme

In certain aspects, the application relates to RNAi, ribozyme, antisense and other nucleic acid-related methods and compositions for manipulating (typically decreasing) a POSH activity. Exemplary RNAi and ribozyme molecules may comprise a sequence as shown in any of SEQ ID Nos: 15, 16, 18, 19, 21, 22, 24 and 25. In certain aspects, the application relates to RNAi, ribozyme, antisense and other nucleic acid-related methods and compositions for manipulating (typically decreasing) a POSH-AP activity. Sequences for examples of UNC48, MSTP028 and HERPUD1 nucleic acids that may be used to design nucleic acids for RNAi, ribozyme, antisense are listed in the Examples.

Certain embodiments of the application make use of materials and methods for effecting knockdown of one or more POSH or POSH-AP genes by means of RNA interference (RNAi). RNAi is a process of sequence-specific post-transcriptional gene repression which can occur in eukaryotic cells. In general, this process involves degradation of an mRNA of a particular sequence induced by double-stranded RNA (dsRNA) that is homologous to that sequence. For example, the expression of a long dsRNA corresponding to the sequence of a particular single-stranded mRNA (ss mRNA) will labilize that message, thereby "interfering" with

expression of the corresponding gene. Accordingly, any selected gene may be repressed by introducing a dsRNA which corresponds to all or a substantial part of the mRNA for that gene. It appears that when a long dsRNA is expressed, it is initially processed by a ribonuclease III into shorter dsRNA oligonucleotides of as few as 21 to 22 base pairs in length. Furthermore, Accordingly, RNAi may be effected by introduction or expression of relatively short homologous dsRNAs. Indeed the use of relatively short homologous dsRNAs may have certain advantages as discussed below.

Mammalian cells have at least two pathways that are affected by double-stranded RNA (dsRNA). In the RNAi (sequence-specific) pathway, the initiating dsRNA is first broken into short interfering (si) RNAs, as described above. The siRNAs have sense and antisense strands of about 21 nucleotides that form approximately 19 nucleotide si RNAs with overhangs of two nucleotides at each 3' end. Short interfering RNAs are thought to provide the sequence information that allows a specific messenger RNA to be targeted for degradation. In contrast, the nonspecific pathway is triggered by dsRNA of any sequence, as long as it is at least about 30 base pairs in length. The nonspecific effects occur because dsRNA activates two enzymes: PKR, which in its active form phosphorylates the translation initiation factor eIF2 to shut down all protein synthesis, and 2', 5' oligoadenylate synthetase (2', 5'-AS), which synthesizes a molecule that activates Rnase L, a nonspecific enzyme that targets all mRNAs. The nonspecific pathway may represents a host response to stress or viral infection, and, in general, the effects of the nonspecific pathway are preferably minimized under preferred methods of the present application. Significantly, longer dsRNAs appear to be required to induce the nonspecific pathway and, accordingly, dsRNAs shorter than about 30 bases pairs are preferred to effect gene repression by RNAi (see Hunter et al. (1975) J Biol Chem 250: 409-17; Manche et al. (1992) Mol Cell Biol 12: 5239-48; Minks et al. (1979) J Biol Chem 254: 10180-3; and Elbashir et al. (2001) Nature 411: 494-8).

RNAi has been shown to be effective in reducing or eliminating the expression of genes in a number of different organisms including *Caenorhabditis elegans* (see e.g. Fire et al. (1998) Nature 391: 806-11), mouse eggs and embryos (Wianny et al. (2000) Nature Cell Biol 2: 70-5; Svoboda et al. (2000) Development

127: 4147-56), and cultured RAT-1 fibroblasts (Bahramina et al. (1999) Mol Cell Biol 19: 274-83), and appears to be an anciently evolved pathway available in eukaryotic plants and animals (Sharp (2001) Genes Dev. 15: 485-90). RNAi has proven to be an effective means of decreasing gene expression in a variety of cell types including HeLa cells, NIH/3T3 cells, COS cells, 293 cells and BHK-21 cells, and typically decreases expression of a gene to lower levels than that achieved using antisense techniques and, indeed, frequently eliminates expression entirely (see Bass (2001) Nature 411: 428-9). In mammalian cells, siRNAs are effective at concentrations that are several orders of magnitude below the concentrations typically used in antisense experiments (Elbashir et al. (2001) Nature 411: 494-8).

The double stranded oligonucleotides used to effect RNAi are preferably less than 30 base pairs in length and, more preferably, comprise about 25, 24, 23, 22, 21, 20, 19, 18 or 17 base pairs of ribonucleic acid. Optionally the dsRNA oligonucleotides of the application may include 3' overhang ends. Exemplary 2-nucleotide 3' overhangs may be composed of ribonucleotide residues of any type and may even be composed of 2'-deoxythymidine residues, which lowers the cost of RNA synthesis and may enhance nuclease resistance of siRNAs in the cell culture medium and within transfected cells (see Elbashir et al. (2001) Nature 411: 494-8). Longer dsRNAs of 50, 75, 100 or even 500 base pairs or more may also be utilized in certain embodiments of the application. Exemplary concentrations of dsRNAs for effecting RNAi are about 0.05 nM, 0.1 nM, 0.5 nM, 1.0 nM, 1.5 nM, 25 nM or 100 nM, although other concentrations may be utilized depending upon the nature of the cells treated, the gene target and other factors readily discernable to the skilled artisan. Exemplary dsRNAs may be synthesized chemically or produced in vitro or in vivo using appropriate expression vectors. Exemplary synthetic RNAs include 21 nucleotide RNAs chemically synthesized using methods known in the art (e.g. Expedite RNA phosphoramidites and thymidine phosphoramidite (Proligo, Germany). Synthetic oligonucleotides are preferably deprotected and gel-purified using methods known in the art (see e.g. Elbashir et al. (2001) Genes Dev. 15: 188-200). Longer RNAs may be transcribed from promoters, such as T7 RNA polymerase promoters, known in the art. A single RNA target, placed in both possible orientations downstream of an in vitro promoter, will transcribe both

strands of the target to create a dsRNA oligonucleotide of the desired target sequence. Any of the above RNA species will be designed to include a portion of nucleic acid sequence represented in a POSH or POSH-AP nucleic acid, such as, for example, a nucleic acid that hybridizes, under stringent and/or physiological conditions, to any of SEQ ID Nos: 1, 3, 4, 6, 8 and 10 and complements thereof or any of the POSH-AP sequences presented in the Examples.

The specific sequence utilized in design of the oligonucleotides may be any contiguous sequence of nucleotides contained within the expressed gene message of the target. Programs and algorithms, known in the art, may be used to select appropriate target sequences. In addition, optimal sequences may be selected utilizing programs designed to predict the secondary structure of a specified single stranded nucleic acid sequence and allowing selection of those sequences likely to occur in exposed single stranded regions of a folded mRNA. Methods and compositions for designing appropriate oligonucleotides may be found, for example, in U.S. Patent Nos. 6,251,588, the contents of which are incorporated herein by reference. Messenger RNA (mRNA) is generally thought of as a linear molecule which contains the information for directing protein synthesis within the sequence of ribonucleotides, however studies have revealed a number of secondary and tertiary structures that exist in most mRNAs. Secondary structure elements in RNA are formed largely by Watson-Crick type interactions between different regions of the same RNA molecule. Important secondary structural elements include intramolecular double stranded regions, hairpin loops, bulges in duplex RNA and internal loops. Tertiary structural elements are formed when secondary structural elements come in contact with each other or with single stranded regions to produce a more complex three dimensional structure. A number of researchers have measured the binding energies of a large number of RNA duplex structures and have derived a set of rules which can be used to predict the secondary structure of RNA (see e.g. Jaeger et al. (1989) Proc. Natl. Acad. Sci. USA 86:7706 (1989); and Turner et al. (1988) Annu. Rev. Biophys. Biophys. Chem. 17:167). The rules are useful in identification of RNA structural elements and, in particular, for identifying single stranded RNA regions which may represent preferred segments of the mRNA to target for silencing RNAi, ribozyme or antisense technologies. Accordingly,

preferred segments of the mRNA target can be identified for design of the RNAi mediating dsRNA oligonucleotides as well as for design of appropriate ribozyme and hammerheadribozyme compositions of the application.

The dsRNA oligonucleotides may be introduced into the cell by transfection
 5 with an heterologous target gene using carrier compositions such as liposomes, which are known in the art- e.g. Lipofectamine 2000 (Life Technologies) as described by the manufacturer for adherent cell lines. Transfection of dsRNA oligonucleotides for targeting endogenous genes may be carried out using
 10 Oligofectamine (Life Technologies). Transfection efficiency may be checked using fluorescence microscopy for mammalian cell lines after co-transfection of hGFP-encoding pAD3 (Kehlenback et al. (1998) J Cell Biol 141: 863-74). The effectiveness of the RNAi may be assessed by any of a number of assays following introduction of the dsRNAs. These include Western blot analysis using antibodies which recognize the POSH or POSH-AP gene product following sufficient time for
 15 turnover of the endogenous pool after new protein synthesis is repressed, reverse transcriptase polymerase chain reaction and Northern blot analysis to determine the level of existing POSH or POSH-AP target mRNA.

Further compositions, methods and applications of RNAi technology are provided in U.S. Patent Application Nos. 6,278,039, 5,723,750 and 5,244,805,
 20 which are incorporated herein by reference.

Ribozyme molecules designed to catalytically cleave POSH or POSH-AP mRNA transcripts can also be used to prevent translation of subject POSH or POSH-AP mRNAs and/or expression of POSH or POSH-APs (see, e.g., PCT International Publication WO90/11364, published October 4, 1990; Sarver et al. (1990) Science
 25 247:1222-1225 and U.S. Patent No. 5,093,246). Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. (For a review, see Rossi (1994) Current Biology 4: 469-471). The mechanism of ribozyme action involves sequence specific hybridization of the ribozyme molecule to complementary target RNA, followed by an endonucleolytic cleavage event. The
 30 composition of ribozyme molecules preferably includes one or more sequences complementary to a POSH or POSH-AP mRNA, and the well known catalytic sequence responsible for mRNA cleavage or a functionally equivalent sequence

(see, e.g., U.S. Pat. No. 5,093,246, which is incorporated herein by reference in its entirety).

While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy target mRNAs, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. Preferably, the target mRNA has the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Haseloff and Gerlach ((1988) *Nature* 334:585-591; and see PCT Appln. No. WO89/05852, the contents of which are incorporated herein by reference). Hammerhead ribozyme sequences can be embedded in a stable RNA such as a transfer RNA (tRNA) to increase cleavage efficiency in vivo (Perriman et al. (1995) *Proc. Natl. Acad. Sci. USA*, 92: 6175-79; de Feyter, and Gaudron, *Methods in Molecular Biology*, Vol. 74, Chapter 43, "Expressing Ribozymes in Plants", Edited by Turner, P. C, Humana Press Inc., Totowa, N.J). In particular, RNA polymerase III-mediated expression of tRNA fusion ribozymes are well known in the art (see Kawasaki et al. (1998) *Nature* 393: 284-9; Kuwabara et al. (1998) *Nature Biotechnol.* 16: 961-5; and Kuwabara et al. (1998) *Mol. Cell* 2: 617-27; Koseki et al. (1999) *J Virol* 73: 1868-77; Kuwabara et al. (1999) *Proc Natl Acad Sci USA* 96: 1886-91; Tanabe et al. (2000) *Nature* 406: 473-4). There are typically a number of potential hammerhead ribozyme cleavage sites within a given target cDNA sequence. Preferably the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the target mRNA- to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts. Furthermore, the use of any cleavage recognition site located in the target sequence encoding different portions of the C-terminal amino acid domains of, for example, long and short forms of target would allow the selective targeting of one or the other form of the target, and thus, have a selective effect on one form of the target gene product.

Gene targeting ribozymes necessarily contain a hybridizing region complementary to two regions, each of at least 5 and preferably each 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 contiguous nucleotides in length of a POSH

or POSH-AP mRNA, such as an mRNA of a sequence represented in any of SEQ ID Nos: 1, 3, 4, 6, 8 or 10 or a POSH-AP presented in the Examples. In addition, ribozymes possess highly specific endoribonuclease activity, which autocatalytically cleaves the target sense mRNA. The present application extends to ribozymes which hybridize to a sense mRNA encoding a POSH gene such as a therapeutic drug target candidate gene, thereby hybridising to the sense mRNA and cleaving it, such that it is no longer capable of being translated to synthesize a functional polypeptide product.

The ribozymes of the present application also include RNA endoribonucleases (hereinafter "Cech-type ribozymes") such as the one which occurs naturally in *Tetrahymena thermophila* (known as the IVS, or L-19 IVS RNA) and which has been extensively described by Thomas Cech and collaborators (Zaug, et al. (1984) *Science* 224:574-578; Zaug, et al. (1986) *Science* 231:470-475; Zaug, et al. (1986) *Nature* 324:429-433; published International patent application No. WO88/04300 by University Patents Inc.; Been, et al. (1986) *Cell* 47:207-216). The Cech-type ribozymes have an eight base pair active site which hybridizes to a target RNA sequence whereafter cleavage of the target RNA takes place. The application encompasses those Cech-type ribozymes which target eight base-pair active site sequences that are present in a target gene or nucleic acid sequence.

Ribozymes can be composed of modified oligonucleotides (e.g., for improved stability, targeting, etc.) and should be delivered to cells which express the target gene in vivo. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive pol III or pol II promoter, so that transfected cells will produce sufficient quantities of the ribozyme to destroy endogenous target messages and inhibit translation. Because ribozymes, unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

In certain embodiments, a ribozyme may be designed by first identifying a sequence portion sufficient to cause effective knockdown by RNAi. The same sequence portion may then be incorporated into a ribozyme. In this aspect of the application, the gene-targeting portions of the ribozyme or RNAi are substantially the same sequence of at least 5 and preferably 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16,

17, 18, 19 or 20 or more contiguous nucleotides of a POSH nucleic acid, such as a nucleic acid of any of SEQ ID Nos: 1, 3, 4, 6, 8, or 10 or POSH-AP nucleic acid, as presented in the Examples. In a long target RNA chain, significant numbers of target sites are not accessible to the ribozyme because they are hidden within secondary or tertiary structures (Birikh et al. (1997) Eur J Biochem 245: 1-16). To overcome the problem of target RNA accessibility, computer generated predictions of secondary structure are typically used to identify targets that are most likely to be single-stranded or have an "open" configuration (see Jaeger et al. (1989) Methods Enzymol 183: 281-306). Other approaches utilize a systematic approach to predicting secondary structure which involves assessing a huge number of candidate hybridizing oligonucleotides molecules (see Milner et al. (1997) Nat Biotechnol 15: 537-41; and Patzel and Sczakiel (1998) Nat Biotechnol 16: 64-8). Additionally, U.S. Patent No. 6,251,588, the contents of which are hereby incorporated herein, describes methods for evaluating oligonucleotide probe sequences so as to predict the potential for hybridization to a target nucleic acid sequence. The method of the application provides for the use of such methods to select preferred segments of a target mRNA sequence that are predicted to be single-stranded and, further, for the opportunistic utilization of the same or substantially identical target mRNA sequence, preferably comprising about 10-20 consecutive nucleotides of the target mRNA, in the design of both the RNAi oligonucleotides and ribozymes of the application.

A further aspect of the application relates to the use of the isolated "antisense" nucleic acids to inhibit expression, e.g., by inhibiting transcription and/or translation of a POSH or POSH-AP nucleic acid. The antisense nucleic acids may bind to the potential drug target by conventional base pair complementarity, or, for example, in the case of binding to DNA duplexes, through specific interactions in the major groove of the double helix. In general, these methods refer to the range of techniques generally employed in the art, and include any methods that rely on specific binding to oligonucleotide sequences.

An antisense construct of the present application can be delivered, for example, as an expression plasmid which, when transcribed in the cell, produces RNA which is complementary to at least a unique portion of the cellular mRNA

which encodes a POSH or POSH-AP polypeptide. Alternatively, the antisense construct is an oligonucleotide probe, which is generated ex vivo and which, when introduced into the cell causes inhibition of expression by hybridizing with the mRNA and/or genomic sequences of a POSH or POSH-AP nucleic acid. Such
 5 oligonucleotide probes are preferably modified oligonucleotides, which are resistant to endogenous nucleases, e.g., exonucleases and/or endonucleases, and are therefore stable in vivo. Exemplary nucleic acid molecules for use as antisense oligonucleotides are phosphoramidate, phosphothioate and methylphosphonate analogs of DNA (see also U.S. Patents 5,176,996; 5,264,564; and 5,256,775).
 10 Additionally, general approaches to constructing oligomers useful in antisense therapy have been reviewed, for example, by Van der Krol et al. (1988) *BioTechniques* 6:958-976; and Stein et al. (1988) *Cancer Res* 48:2659- 2668.

With respect to antisense DNA, oligodeoxyribonucleotides derived from the translation initiation site, e.g., between the -10 and +10 regions of the target gene,
 15 are preferred. Antisense approaches involve the design of oligonucleotides (either DNA or RNA) that are complementary to mRNA encoding a POSH or POSH-AP polypeptide. The antisense oligonucleotides will bind to the mRNA transcripts and prevent translation. Absolute complementarity, although preferred, is not required. In the case of double-stranded antisense nucleic acids, a single strand of the duplex
 20 DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with an RNA it may contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of
 25 mismatch by use of standard procedures to determine the melting point of the hybridized complex.

Oligonucleotides that are complementary to the 5' end of the mRNA, e.g., the 5' untranslated sequence up to and including the AUG initiation codon, should work most efficiently at inhibiting translation. However, sequences complementary to the
 30 3' untranslated sequences of mRNAs have recently been shown to be effective at inhibiting translation of mRNAs as well. (Wagner, R. 1994. *Nature* 372:333). Therefore, oligonucleotides complementary to either the 5' or 3' untranslated, non-

coding regions of a gene could be used in an antisense approach to inhibit translation of that mRNA. Oligonucleotides complementary to the 5' untranslated region of the mRNA should include the complement of the AUG start codon. Antisense oligonucleotides complementary to mRNA coding regions are less efficient inhibitors of translation but could also be used in accordance with the application. Whether designed to hybridize to the 5', 3' or coding region of mRNA, antisense nucleic acids should be at least six nucleotides in length, and are preferably less than about 100 and more preferably less than about 50, 25, 17 or 10 nucleotides in length.

It is preferred that in vitro studies are first performed to quantitate the ability of the antisense oligonucleotide to inhibit gene expression. It is preferred that these studies utilize controls that distinguish between antisense gene inhibition and nonspecific biological effects of oligonucleotides. It is also preferred that these studies compare levels of the target RNA or protein with that of an internal control RNA or protein. Results obtained using the antisense oligonucleotide may be compared with those obtained using a control oligonucleotide. It is preferred that the control oligonucleotide is of approximately the same length as the test oligonucleotide and that the nucleotide sequence of the oligonucleotide differs from the antisense sequence no more than is necessary to prevent specific hybridization to the target sequence.

The antisense oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556; Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No. W088/09810, published December 15, 1988) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134, published April 25, 1988), hybridization-triggered cleavage agents. (See, e.g., Krol et al., 1988, BioTechniques 6:958-976) or intercalating agents. (See, e.g., Zon, 1988, Pharm. Res. 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide,

hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The antisense oligonucleotide may comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5- bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4- acetylcytosine, 5- (carboxyhydroxyethyl) uracil, 5-carboxymethylaminomethyl-2- thiouridine, 5- carboxymethylaminomethyluracil, dihydrouracil, beta-D- galactosylqueosine, inosine, N6- isopentenyladenine, 1-methylguanine, 1- methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3- methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5- methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D- mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio- N6- isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5- methyluracil, uracil-5- oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5- methyl-2-thiouracil, 3-(3-amino-3- N-2-carboxypropyl) uracil, (acp3)w, and 2,6- diaminopurine.

The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2- fluoroarabinose, xylulose, and hexose.

The antisense oligonucleotide can also contain a neutral peptide-like backbone. Such molecules are termed peptide nucleic acid (PNA)-oligomers and are described, e.g., in Perry-O'Keefe et al. (1996) Proc. Natl. Acad. Sci. U.S.A. 93:14670 and in Eglom et al. (1993) Nature 365:566. One advantage of PNA oligomers is their capability to bind to complementary DNA essentially independently from the ionic strength of the medium due to the neutral backbone of the DNA. In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In yet a further embodiment, the antisense oligonucleotide is an alpha-anomeric oligonucleotide. An alpha-anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual antiparallel orientation, the strands run parallel to each other (Gautier et al., 1987, Nucl. Acids Res. 15:6625-6641). The oligonucleotide is a 2'-O-methylribonucleotide (Inoue et al., 1987, Nucl. Acids Res. 15:6131-6148), or a chimeric RNA-DNA analogue (Inoue et al., 1987, FEBS Lett. 215:327-330).

While antisense nucleotides complementary to the coding region of a POSH or POSH-AP mRNA sequence can be used, those complementary to the transcribed untranslated region may also be used.

In certain instances, it may be difficult to achieve intracellular concentrations of the antisense sufficient to suppress translation on endogenous mRNAs. Therefore a preferred approach utilizes a recombinant DNA construct in which the antisense oligonucleotide is placed under the control of a strong pol III or pol II promoter. The use of such a construct to transfect target cells will result in the transcription of sufficient amounts of single stranded RNAs that will form complementary base pairs with the endogenous potential drug target transcripts and thereby prevent translation. For example, a vector can be introduced such that it is taken up by a cell and directs the transcription of an antisense RNA. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in mammalian cells. Expression of the sequence encoding the antisense RNA can be by any promoter known in the art to act in mammalian, preferably human cells. Such promoters can be inducible or constitutive. Such promoters include but are not limited to: the SV40 early promoter region (Bernoist and Chambon, 1981, Nature 290:304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., 1980, Cell 22:787-797), the herpes thymidine kinase promoter (Wagner et al., 1981, Proc. Natl. Acad. Sci. U.S.A. 78:1441-1445), the regulatory sequences of the metallothionein gene (Brinster et al, 1982, Nature 296:39-42), etc.

Any type of plasmid, cosmid, YAC or viral vector can be used to prepare the recombinant DNA construct, which can be introduced directly into the tissue site.

Alternatively, POSH or POSH-AP gene expression can be reduced by targeting deoxyribonucleotide sequences complementary to the regulatory region of the gene (i.e., the promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells in the body. (See generally, Helene, C. 1991, *Anticancer Drug Des.*, 6(6):569-84; Helene, C., et al., 1992, *Ann. N.Y. Acad. Sci.*, 660:27-36; and Maher, L.J., 1992, *Bioassays* 14(12):807-15).

Nucleic acid molecules to be used in triple helix formation for the inhibition of transcription are preferably single stranded and composed of deoxyribonucleotides. The base composition of these oligonucleotides should promote triple helix formation via Hoogsteen base pairing rules, which generally require sizable stretches of either purines or pyrimidines to be present on one strand of a duplex. Nucleotide sequences may be pyrimidine-based, which will result in TAT and CGC triplets across the three associated strands of the resulting triple helix. The pyrimidine-rich molecules provide base complementarity to a purine-rich region of a single strand of the duplex in a parallel orientation to that strand. In addition, nucleic acid molecules may be chosen that are purine-rich, for example, containing a stretch of G residues. These molecules will form a triple helix with a DNA duplex that is rich in GC pairs, in which the majority of the purine residues are located on a single strand of the targeted duplex, resulting in CGC triplets across the three strands in the triplex.

Alternatively, POSH or POSH-AP sequences that can be targeted for triple helix formation may be increased by creating a so called "switchback" nucleic acid molecule. Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizable stretch of either purines or pyrimidines to be present on one strand of a duplex.

A further aspect of the application relates to the use of DNA enzymes to inhibit expression of a POSH or POSH-AP gene. DNA enzymes incorporate some of the mechanistic features of both antisense and ribozyme technologies. DNA enzymes are designed so that they recognize a particular target nucleic acid

sequence, much like an antisense oligonucleotide, however much like a ribozyme they are catalytic and specifically cleave the target nucleic acid.

There are currently two basic types of DNA enzymes, and both of these were identified by Santoro and Joyce (see, for example, US Patent No. 6110462). The
5 10-23 DNA enzyme comprises a loop structure which connect two arms. The two arms provide specificity by recognizing the particular target nucleic acid sequence while the loop structure provides catalytic function under physiological conditions.

Briefly, to design an ideal DNA enzyme that specifically recognizes and cleaves a target nucleic acid, one of skill in the art must first identify the unique
10 target sequence. This can be done using the same approach as outlined for antisense oligonucleotides. Preferably, the unique or substantially sequence is a G/C rich of approximately 18 to 22 nucleotides. High G/C content helps insure a stronger interaction between the DNA enzyme and the target sequence.

When synthesizing the DNA enzyme, the specific antisense recognition
15 sequence that will target the enzyme to the message is divided so that it comprises the two arms of the DNA enzyme, and the DNA enzyme loop is placed between the two specific arms.

Methods of making and administering DNA enzymes can be found, for example, in US 6110462. Similarly, methods of delivery DNA ribozymes in vitro
20 or in vivo include methods of delivery RNA ribozyme, as outlined in detail above. Additionally, one of skill in the art will recognize that, like antisense oligonucleotide, DNA enzymes can be optionally modified to improve stability and improve resistance to degradation.

Antisense RNA and DNA, ribozyme, RNAi and triple helix molecules of the
25 application may be prepared by any method known in the art for the synthesis of DNA and RNA molecules. These include techniques for chemically synthesizing oligodeoxyribonucleotides and oligoribonucleotides well known in the art such as for example solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences
30 encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors which incorporate suitable RNA polymerase

promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines. Moreover, various well-known modifications to nucleic acid molecules may be introduced as a means of increasing intracellular stability and half-life. Possible modifications include but are not limited to the addition of flanking sequences of ribonucleotides or deoxyribonucleotides to the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the oligodeoxyribonucleotide backbone.

9. Drug Screening Assays

In certain aspects, the present application also provides assays for identifying therapeutic agents which either interfere with or promote POSH or POSH-AP function. In certain embodiments, agents of the application are antiviral agents, optionally interfering with viral maturation, and preferably where the virus is a retroid virus, an RNA virus and an envelope virus. In certain preferred embodiments, an antiviral agent interferes with the ubiquitin ligase catalytic activity of POSH (e.g. POSH auto-ubiquitination or transfer to a target protein). In certain preferred embodiments, an antiviral agent interferes with the interaction between POSH and a POSH-AP polypeptide, for example an antiviral agent may disrupt or render irreversible the interaction between a POSH polypeptide and POSH-AP polypeptide such as an UNC84, an MSTP028, a HERPUD1, another POSH polypeptide (as in the case of a POSH dimer, a heterodimer of two different POSH polypeptides, homomultimers and heteromultimers); a GTPase (eg. Rac, Rac1, Rho, Ras); an E2 enzyme and ubiquitin, or optionally, a cullin; a clathrin; AP-1; AP-2; an HSP70; an HSP90, Brca1, Bard1, Nef, PAK1, PAK2, PAK family, Vav, Cdc42, PI3K (e.g. p85 or p110), Nedd4, src (src family), a Gag, particularly an HIV Gag (e.g. p160), Tsg101, VASP, RNB6, WASP, N-WASP and KIAA0674, Similar to Spred-2, as well as, in certain embodiments, proteins known to be associated with clathrin-coated vesicles and or proteins involved in the protein sorting pathway. In further embodiments, agents of the application are anti-apoptotic agents, optionally interfering with JNK and/or NF- κ B signaling. In yet additional embodiments,

agents of the application interfere with the signaling of a GTPase, such as Rac or Ras, optionally disrupting the interaction between a POSH polypeptide and a Rac protein. In certain embodiments, agents of the application modulate the ubiquitin ligase activity of POSH and may be used to treat certain diseases related to ubiquitin
 5 ligase activity.

In certain embodiments, the application provides assays to identify, optimize or otherwise assess agents that increase or decrease a ubiquitin-related activity of a POSH polypeptide. Ubiquitin-related activities of POSH polypeptides may include the self-ubiquitination activity of a POSH polypeptide, generally involving the
 10 transfer of ubiquitin from an E2 enzyme to the POSH polypeptide, and the ubiquitination of a target protein, generally involving the transfer of a ubiquitin from a POSH polypeptide to the target protein. In certain embodiments, a POSH activity is mediated, at least in part, by a POSH RING domain.

In certain embodiments, an assay comprises forming a mixture comprising a
 15 POSH polypeptide, an E2 polypeptide and a source of ubiquitin (which may be the E2 polypeptide pre-complexed with ubiquitin). Optionally the mixture comprises an E1 polypeptide and optionally the mixture comprises a target polypeptide. Additional components of the mixture may be selected to provide conditions consistent with the ubiquitination of the POSH polypeptide. One or more of a
 20 variety of parameters may be detected, such as POSH-ubiquitin conjugates, E2-ubiquitin thioesters, free ubiquitin and target polypeptide-ubiquitin complexes. The term "detect" is used herein to include a determination of the presence or absence of the subject of detection (e.g. POSH-ubiquitin, E2-ubiquitin, etc.), a quantitative measure of the amount of the subject of detection, or a mathematical calculation of
 25 the presence, absence or amount of the subject of detection, based on the detection of other parameters. The term "detect" includes the situation wherein the subject of detection is determined to be absent or below the level of sensitivity. Detection may comprise detection of a label (e.g. fluorescent label, radioisotope label, and other described below), resolution and identification by size (e.g. SDS-PAGE, mass
 30 spectroscopy), purification and detection, and other methods that, in view of this specification, will be available to one of skill in the art. For instance, radioisotope labeling may be measured by scintillation counting, or by densitometry after

exposure to a photographic emulsion, or by using a device such as a Phosphorimager. Likewise, densitometry may be used to measure bound ubiquitin following a reaction with an enzyme label substrate that produces an opaque product when an enzyme label is used. In a preferred embodiment, an assay comprises
5 detecting the POSH-ubiquitin conjugate.

In certain embodiments, an assay comprises forming a mixture comprising a POSH polypeptide, a target polypeptide and a source of ubiquitin (which may be the POSH polypeptide pre-complexed with ubiquitin). Optionally the mixture comprises an E1 and/or E2 polypeptide and optionally the mixture comprises an E2-
10 ubiquitin thioester. Additional components of the mixture may be selected to provide conditions consistent with the ubiquitination of the target polypeptide. One or more of a variety of parameters may be detected, such as POSH-ubiquitin conjugates and target polypeptide-ubiquitin conjugates. In a preferred embodiment, an assay comprises detecting the target polypeptide-ubiquitin conjugate. In another
15 preferred embodiment, an assay comprises detecting the POSH-ubiquitin conjugate.

An assay described above may be used in a screening assay to identify agents that modulate a ubiquitin-related activity of a POSH polypeptide. A screening assay will generally involve adding a test agent to one of the above assays, or any other assay designed to assess a ubiquitin-related activity of a POSH polypeptide. The
20 parameter(s) detected in a screening assay may be compared to a suitable reference. A suitable reference may be an assay run previously, in parallel or later that omits the test agent. A suitable reference may also be an average of previous measurements in the absence of the test agent. In general the components of a screening assay mixture may be added in any order consistent with the overall
25 activity to be assessed, but certain variations may be preferred. For example, in certain embodiments, it may be desirable to pre-incubate the test agent and the E3 (e.g. the POSH polypeptide), followed by removing the test agent and addition of other components to complete the assay. In this manner, the effects of the agent solely on the POSH polypeptide may be assessed. In certain preferred
30 embodiments, a screening assay for an antiviral agent employs a target polypeptide comprising an L domain, and preferably an HIV L domain.

In certain embodiments, an assay is performed in a high-throughput format. For example, one of the components of a mixture may be affixed to a solid substrate and one or more of the other components is labeled. For example, the POSH polypeptide may be affixed to a surface, such as a 96-well plate, and the ubiquitin is in solution and labeled. An E2 and E1 are also in solution, and the POSH-ubiquitin conjugate formation may be measured by washing the solid surface to remove uncomplexed labeled ubiquitin and detecting the ubiquitin that remains bound. Other variations may be used. For example, the amount of ubiquitin in solution may be detected. In certain embodiments, the formation of ubiquitin complexes may be measured by an interactive technique, such as FRET, wherein a ubiquitin is labeled with a first label and the desired complex partner (e.g. POSH polypeptide or target polypeptide) is labeled with a second label, wherein the first and second label interact when they come into close proximity to produce an altered signal. In FRET, the first and second labels are fluorophores. FRET is described in greater detail below. The formation of polyubiquitin complexes may be performed by mixing two or more pools of differentially labeled ubiquitin that interact upon formation of a polyubiquitin (see, e.g. US Patent Publication 20020042083). High-throughput may be achieved by performing an interactive assay, such as FRET, in solution as well. In addition, if a polypeptide in the mixture, such as the POSH polypeptide or target polypeptide, is readily purifiable (e.g. with a specific antibody or via a tag such as biotin, FLAG, polyhistidine, etc.), the reaction may be performed in solution and the tagged polypeptide rapidly isolated, along with any polypeptides, such as ubiquitin, that are associated with the tagged polypeptide. Proteins may also be resolved by SDS-PAGE for detection.

In certain embodiments, the ubiquitin is labeled, either directly or indirectly. This typically allows for easy and rapid detection and measurement of ligated ubiquitin, making the assay useful for high-throughput screening applications. As described above, certain embodiments may employ one or more tagged or labeled proteins. A "tag" is meant to include moieties that facilitate rapid isolation of the tagged polypeptide. A tag may be used to facilitate attachment of a polypeptide to a surface. A "label" is meant to include moieties that facilitate rapid detection of the labeled polypeptide. Certain moieties may be used both as a label and a tag (e.g.

epitope tags that are readily purified and detected with a well-characterized antibody). Biotinylation of polypeptides is well known, for example, a large number of biotinylation agents are known, including amine-reactive and thiol-reactive agents, for the biotinylation of proteins, nucleic acids, carbohydrates, carboxylic acids; see chapter 4, Molecular Probes Catalog, Haugland, 6th Ed. 1996, hereby
 5 incorporated by reference. A biotinylated substrate can be attached to a biotinylated component via avidin or streptavidin. Similarly, a large number of haptenylation reagents are also known.

An "E1" is a ubiquitin activating enzyme. In a preferred embodiment, E1 is
 10 capable of transferring ubiquitin to an E2. In a preferred embodiment, E1 forms a high energy thiolester bond with ubiquitin, thereby "activating" the ubiquitin. An "E2" is a ubiquitin carrier enzyme (also known as a ubiquitin conjugating enzyme). In a preferred embodiment, ubiquitin is transferred from E1 to E2. In a preferred embodiment, the transfer results in a thiolester bond formed between E2 and
 15 ubiquitin. In a preferred embodiment, E2 is capable of transferring ubiquitin to a POSH polypeptide.

In an alternative embodiment, a POSH polypeptide, E2 or target polypeptide is bound to a bead, optionally with the assistance of a tag. Following ligation, the beads may be separated from the unbound ubiquitin and the bound ubiquitin
 20 measured. In a preferred embodiment, POSH polypeptide is bound to beads and the composition used includes labeled ubiquitin. In this embodiment, the beads with bound ubiquitin may be separated using a fluorescence-activated cell sorting (FACS) machine. Methods for such use are described in U.S. patent application Ser. No. 09/047,119, which is hereby incorporated in its entirety. The amount of bound
 25 ubiquitin can then be measured.

In a screening assay, the effect of a test agent may be assessed by, for example, assessing the effect of the test agent on kinetics, steady-state and/or endpoint of the reaction.

The components of the various assay mixtures provided herein may be
 30 combined in varying amounts. In a preferred embodiment, ubiquitin (or E2

complexed ubiquitin) is combined at a final concentration of from 5 to 200 ng per 100 microliter reaction solution. Optionally E1 is used at a final concentration of from 1 to 50 ng per 100 microliter reaction solution. Optionally E2 is combined at a final concentration of 10 to 100 ng per 100 microliter reaction solution, more preferably 10-50 ng per 100 microliter reaction solution. In a preferred embodiment, POSH polypeptide is combined at a final concentration of from 1 ng to 500 ng per 100 microliter reaction solution.

Generally, an assay mixture is prepared so as to favor ubiquitin ligase activity and/or ubiquitination activity. Generally, this will be physiological conditions, such as 50 – 200 mM salt (e.g. NaCl, KCl), pH of between 5 and 9, and preferably between 6 and 8. Such conditions may be optimized through trial and error. Incubations may be performed at any temperature which facilitates optimal activity, typically between 4 and 40 degrees C. Incubation periods are selected for optimum activity, but may also be optimized to facilitate rapid high throughput screening. Typically between 0.5 and 1.5 hours will be sufficient. A variety of other reagents may be included in the compositions. These include reagents like salts, solvents, buffers, neutral proteins, e.g. albumin, detergents, etc. which may be used to facilitate optimal ubiquitination enzyme activity and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The compositions will also preferably include adenosine tri-phosphate (ATP). The mixture of components may be added in any order that promotes ubiquitin ligase activity or optimizes identification of candidate modulator effects. In a preferred embodiment, ubiquitin is provided in a reaction buffer solution, followed by addition of the ubiquitination enzymes. In an alternate preferred embodiment, ubiquitin is provided in a reaction buffer solution, a candidate modulator is then added, followed by addition of the ubiquitination enzymes.

In general, a test agent that decreases a POSH ubiquitin-related activity may be used to inhibit POSH function in vivo, while a test agent that increases a POSH ubiquitin-related activity may be used to stimulate POSH function in vivo. Test

agent may be modified for use in vivo, e.g. by addition of a hydrophobic moiety, such as an ester.

An additional POSH-AP may be added to a POSH ubiquitination assay to assess the effect of the POSH-AP on POSH-mediated ubiquitination and/or to assess whether the POSH-AP is a target for POSH-mediated ubiquitination.

Certain embodiments of the application relate to assays for identifying agents that bind to a POSH or POSH-AP polypeptide, optionally a particular domain of POSH such as an SH3 or RING domain or a particular domain of a POSH-AP, such as a BTB/POZ domain of an MSTP028 or a ubiquitin-like domain of a HERPUD1.

In preferred embodiments, a POSH polypeptide is a polypeptide comprising the fourth SH3 domain of hPOSH (SEQ ID NO: 30). A wide variety of assays may be used for this purpose, including labeled in vitro protein-protein binding assays, electrophoretic mobility shift assays, immunoassays for protein binding, and the like. The purified protein may also be used for determination of three-dimensional crystal structure, which can be used for modeling intermolecular interactions and design of test agents. In one embodiment, an assay detects agents which inhibit interaction of one or more subject POSH polypeptides with a POSH-AP. In another embodiment, the assay detects agents which modulate the intrinsic biological activity of a POSH polypeptide or POSH complex, such as an enzymatic activity, binding to other cellular components, cellular compartmentalization, and the like.

In one aspect, the application provides methods and compositions for the identification of compositions that interfere with the function of POSH or POSH-AP polypeptides. Given the role of POSH polypeptides in viral production, compositions that perturb the formation or stability of the protein-protein interactions between POSH polypeptides and the proteins that they interact with, such as POSH-APs, and particularly POSH complexes comprising a viral protein, are candidate pharmaceuticals for the treatment of viral infections.

While not wishing to be bound to mechanism, it is postulated that POSH polypeptides promote the assembly of protein complexes that are important in release of virions and other biological processes. Complexes of the application may include a combination of a POSH polypeptide and a POSH-AP. Exemplary

complexes may comprise one or more of the following: a POSH polypeptide (as in the case of a POSH dimer, a heterodimer of two different POSH, homomultimers and heteromultimers); an UNC48, an MSTP028, an HERPUD1, a GTPase (eg. Rac, Rac1, Rho, Ras); an E2 enzyme; ubiquitin, or optionally, a cullin; a clathrin; AP-1; 5 AP-2; an HSP70; an HSP90, Brcal, Bard1, Nef, PAK1, PAK2, PAK family, Vav, Cdc42, PI3K (e.g. p85 or p110), Nedd4, src (src family), Tsg101, VASP, RNB6, WASP, N-WASP, a Gag, particularly an HIV Gag (e.g. p160); and KIAA0674, Similar to Spred-2, as well as, in certain embodiments, proteins known to be associated with clathrin-coated vesicles and or proteins involved in the protein 10 sorting pathway.

The type of complex formed by a POSH polypeptide will depend upon the domains present in the protein. While not intended to be limiting, exemplary domains of potential interacting proteins are provided below. A RING domain is expected to interact with cullins, E2 enzymes, AP-1, AP-2, and/or a substrate for 15 ubiquitylation (e.g. in some instances, a protein comprising a Gag L domain or a Gag polypeptide such as Gag-Pol, such as HIV p160). An SH3 domain may interact with Gag L domains and other proteins having the sequence motif P(T/S)AP, RXXP(T/S)AP, PXXDY, PXXP, PPXY or RXXPPXXP, such as, for example, an HIV Gag sequence such as RQGPKPEFR, PFRDY, PTAP and RPEPTAP.

20 In a preferred assay for an antiviral or antiapoptotic agent, the test agent is assessed for its ability to disrupt or inhibit the formation of a complex of a POSH polypeptide and a Rac polypeptide, particularly a human Rac polypeptide, such as Rac1.

A variety of assay formats will suffice and, in light of the present disclosure, 25 those not expressly described herein will nevertheless be comprehended by one of ordinary skill in the art. Assay formats which approximate such conditions as formation of protein complexes, enzymatic activity, and even a POSH polypeptide-mediated membrane reorganization or vesicle formation activity, may be generated in many different forms, and include assays based on cell-free systems, e.g. purified 30 proteins or cell lysates, as well as cell-based assays which utilize intact cells. Simple binding assays can also be used to detect agents which bind to POSH. Such binding assays may also identify agents that act by disrupting the interaction

between a POSH polypeptide and a POSH interacting protein, or the binding of a POSH polypeptide or complex to a substrate. Agents to be tested can be produced, for example, by bacteria, yeast or other organisms (e.g. natural products), produced chemically (e.g. small molecules, including peptidomimetics), or produced
5 recombinantly. In a preferred embodiment, the test agent is a small organic molecule, e.g., other than a peptide or oligonucleotide, having a molecular weight of less than about 2,000 daltons.

In many drug screening programs which test libraries of compounds and natural extracts, high throughput assays are desirable in order to maximize the
10 number of compounds surveyed in a given period of time. Assays of the present application which are performed in cell-free systems, such as may be developed with purified or semi-purified proteins or with lysates, are often preferred as "primary" screens in that they can be generated to permit rapid development and relatively easy detection of an alteration in a molecular target which is mediated by a test
15 compound. Moreover, the effects of cellular toxicity and/or bioavailability of the test compound can be generally ignored in the in vitro system, the assay instead being focused primarily on the effect of the drug on the molecular target as may be manifest in an alteration of binding affinity with other proteins or changes in enzymatic properties of the molecular target.

20 In preferred in vitro embodiments of the present assay, a reconstituted POSH complex comprises a reconstituted mixture of at least semi-purified proteins. By semi-purified, it is meant that the proteins utilized in the reconstituted mixture have been previously separated from other cellular or viral proteins. For instance, in contrast to cell lysates, the proteins involved in POSH complex formation are
25 present in the mixture to at least 50% purity relative to all other proteins in the mixture, and more preferably are present at 90-95% purity. In certain embodiments of the subject method, the reconstituted protein mixture is derived by mixing highly purified proteins such that the reconstituted mixture substantially lacks other proteins (such as of cellular or viral origin) which might interfere with or otherwise
30 alter the ability to measure POSH complex assembly and/or disassembly.

Assaying POSH complexes, in the presence and absence of a candidate inhibitor, can be accomplished in any vessel suitable for containing the reactants. Examples include microtitre plates, test tubes, and micro-centrifuge tubes.

5 In one embodiment of the present application, drug screening assays can be generated which detect inhibitory agents on the basis of their ability to interfere with assembly or stability of the POSH complex. In an exemplary binding assay, the compound of interest is contacted with a mixture comprising a POSH polypeptide and at least one interacting polypeptide. Detection and quantification of POSH
10 (or potentiating) interaction between the two polypeptides. The efficacy of the compound can be assessed by generating dose response curves from data obtained using various concentrations of the test compound. Moreover, a control assay can also be performed to provide a baseline for comparison. In the control assay, the formation of complexes is quantitated in the absence of the test compound.

15 Complex formation between the POSH polypeptides and a substrate polypeptide may be detected by a variety of techniques, many of which are effectively described above. For instance, modulation in the formation of complexes can be quantitated using, for example, detectably labeled proteins (e.g. radiolabeled, fluorescently labeled, or enzymatically labeled), by immunoassay, or by
20 chromatographic detection. Surface plasmon resonance systems, such as those available from Biacore International AB (Uppsala, Sweden), may also be used to detect protein-protein interaction

Often, it will be desirable to immobilize one of the polypeptides to facilitate separation of complexes from uncomplexed forms of one of the proteins, as well as
25 to accommodate automation of the assay. In an illustrative embodiment, a fusion protein can be provided which adds a domain that permits the protein to be bound to an insoluble matrix. For example, GST-POSH fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with a potential interacting
30 protein, e.g. an ³⁵S-labeled polypeptide, and the test compound and incubated under conditions conducive to complex formation. Following incubation, the beads are washed to remove any unbound interacting protein, and the matrix bead-bound

radiolabel determined directly (e.g. beads placed in scintillant), or in the supernatant after the complexes are dissociated, e.g. when microtitre plate is used. Alternatively, after washing away unbound protein, the complexes can be dissociated from the matrix, separated by SDS-PAGE gel, and the level of interacting polypeptide found
 5 in the matrix-bound fraction quantitated from the gel using standard electrophoretic techniques.

In a further embodiment, agents that bind to a POSH or POSH-AP may be identified by using an immobilized POSH or POSH-AP. In an illustrative embodiment, a fusion protein can be provided which adds a domain that permits the
 10 protein to be bound to an insoluble matrix. For example, GST-POSH fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with a potential labeled binding agent and incubated under conditions conducive to binding. Following incubation, the beads are washed to remove any unbound agent,
 15 and the matrix bead-bound label determined directly, or in the supernatant after the bound agent is dissociated.

In yet another embodiment, the POSH polypeptide and potential interacting polypeptide can be used to generate an interaction trap assay (see also, U.S. Patent NO: 5,283,317; Zervos et al. (1993) Cell 72:223-232; Madura et al. (1993) J Biol
 20 Chem 268:12046-12054; Bartel et al. (1993) Biotechniques 14:920-924; and Iwabuchi et al. (1993) Oncogene 8:1693-1696), for subsequently detecting agents which disrupt binding of the proteins to one and other.

In particular, the method makes use of chimeric genes which express hybrid proteins. To illustrate, a first hybrid gene comprises the coding sequence for a
 25 DNA-binding domain of a transcriptional activator can be fused in frame to the coding sequence for a "bait" protein, e.g., a POSH polypeptide of sufficient length to bind to a potential interacting protein. The second hybrid protein encodes a transcriptional activation domain fused in frame to a gene encoding a "fish" protein, e.g., a potential interacting protein of sufficient length to interact with the POSH
 30 polypeptide portion of the bait fusion protein. If the bait and fish proteins are able to interact, e.g., form a POSH complex, they bring into close proximity the two domains of the transcriptional activator. This proximity causes transcription of a

reporter gene which is operably linked to a transcriptional regulatory site responsive to the transcriptional activator, and expression of the reporter gene can be detected and used to score for the interaction of the bait and fish proteins.

In accordance with the present application, the method includes providing a
 5 host cell, preferably a yeast cell, e.g., *Kluyveri lactis*, *Schizosaccharomyces pombe*, *Ustilago maydis*, *Saccharomyces cerevisiae*, *Neurospora crassa*, *Aspergillus niger*, *Aspergillus nidulans*, *Pichia pastoris*, *Candida tropicalis*, and *Hansenula polymorpha*, though most preferably *S. cerevisiae* or *S. pombe*. The host cell contains a reporter gene having a binding site for the DNA-binding domain of a
 10 transcriptional activator used in the bait protein, such that the reporter gene expresses a detectable gene product when the gene is transcriptionally activated. The first chimeric gene may be present in a chromosome of the host cell, or as part of an expression vector. Interaction trap assays may also be performed in mammalian and bacterial cell types.

15 The host cell also contains a first chimeric gene which is capable of being expressed in the host cell. The gene encodes a chimeric protein, which comprises (i) a DNA-binding domain that recognizes the responsive element on the reporter gene in the host cell, and (ii) a bait protein, such as a POSH polypeptide sequence.

A second chimeric gene is also provided which is capable of being expressed
 20 in the host cell, and encodes the "fish" fusion protein. In one embodiment, both the first and the second chimeric genes are introduced into the host cell in the form of plasmids. Preferably, however, the first chimeric gene is present in a chromosome of the host cell and the second chimeric gene is introduced into the host cell as part of a plasmid.

25 Preferably, the DNA-binding domain of the first hybrid protein and the transcriptional activation domain of the second hybrid protein are derived from transcriptional activators having separable DNA-binding and transcriptional activation domains. For instance, these separate DNA-binding and transcriptional activation domains are known to be found in the yeast GAL4 protein, and are known
 30 to be found in the yeast GCN4 and ADR1 proteins. Many other proteins involved in transcription also have separable binding and transcriptional activation domains which make them useful for the present application, and include, for example, the

LexA and VP16 proteins. It will be understood that other (substantially) transcriptionally-inert DNA-binding domains may be used in the subject constructs; such as domains of ACE1, Icl, lac repressor, jun or fos. In another embodiment, the DNA-binding domain and the transcriptional activation domain may be from
5 different proteins. The use of a LexA DNA binding domain provides certain advantages. For example, in yeast, the LexA moiety contains no activation function and has no known effect on transcription of yeast genes. In addition, use of LexA allows control over the sensitivity of the assay to the level of interaction (see, for example, the Brent et al. PCT publication WO94/10300).

10 In preferred embodiments, any enzymatic activity associated with the bait or fish proteins is inactivated, e.g., dominant negative or other mutants of a POSH polypeptide can be used.

Continuing with the illustrated example, the POSH polypeptide-mediated interaction, if any, between the bait and fish fusion proteins in the host cell,
15 therefore, causes the activation domain to activate transcription of the reporter gene. The method is carried out by introducing the first chimeric gene and the second chimeric gene into the host cell, and subjecting that cell to conditions under which the bait and fish fusion proteins are expressed in sufficient quantity for the reporter gene to be activated. The formation of a POSH - POSH-AP complex results
20 in a detectable signal produced by the expression of the reporter gene. Accordingly, the level of formation of a complex in the presence of a test compound and in the absence of the test compound can be evaluated by detecting the level of expression of the reporter gene in each case. Various reporter constructs may be used in accord with the methods of the application and include, for example, reporter genes which
25 produce such detectable signals as selected from the group consisting of an enzymatic signal, a fluorescent signal, a phosphorescent signal and drug resistance.

One aspect of the present application provides reconstituted protein preparations including a POSH polypeptide and one or more interacting polypeptides.

30 In still further embodiments of the present assay, the POSH complex is generated in whole cells, taking advantage of cell culture techniques to support the subject assay. For example, as described below, the POSH complex can be

constituted in a eukaryotic cell culture system, including mammalian and yeast cells. Often it will be desirable to express one or more viral proteins (eg. Gag or Env) in such a cell along with a subject POSH polypeptide. It may also be desirable to infect the cell with a virus of interest. Advantages to generating the subject assay in
 5 an intact cell include the ability to detect inhibitors which are functional in an environment more closely approximating that which therapeutic use of the inhibitor would require, including the ability of the agent to gain entry into the cell. Furthermore, certain of the in vivo embodiments of the assay, such as examples given below, are amenable to high through-put analysis of candidate agents.

10 The components of the POSH complex can be endogenous to the cell selected to support the assay. Alternatively, some or all of the components can be derived from exogenous sources. For instance, fusion proteins can be introduced into the cell by recombinant techniques (such as through the use of an expression vector), as well as by microinjecting the fusion protein itself or mRNA encoding the
 15 fusion protein.

In many embodiments, a cell is manipulated after incubation with a candidate agent and assayed for a POSH or POSH-AP activity. In certain embodiments a POSH or POSH-AP activity is represented by production of virus like particles. As demonstrated herein, an agent that disrupts POSH or POSH-AP
 20 activity can cause a decrease in the production of virus like particles. Other bioassays for POSH or POSH-AP activities may include apoptosis assays (e.g. cell survival assays, apoptosis reporter gene assays, etc.) and NF-kB nuclear localization assays (see e.g. Tapon et al. (1998) EMBO J. 17: 1395-1404). In certain
 25 embodiments, POSH or POSH-AP activities may include, without limitation, complex formation, ubiquitination and membrane fusion events (eg. release of viral buds or fusion of vesicles). POSH complex formation may be assessed by immunoprecipitation and analysis of co-immunoprecipitated proteins or affinity purification and analysis of co-purified proteins. Fluorescence Resonance Energy Transfer (FRET)-based assays may also be used to determine complex formation.
 30 Fluorescent molecules having the proper emission and excitation spectra that are brought into close proximity with one another can exhibit FRET. The fluorescent molecules are chosen such that the emission spectrum of one of the molecules (the

donor molecule) overlaps with the excitation spectrum of the other molecule (the acceptor molecule). The donor molecule is excited by light of appropriate intensity within the donor's excitation spectrum. The donor then emits the absorbed energy as fluorescent light. The fluorescent energy it produces is quenched by the acceptor molecule. FRET can be manifested as a reduction in the intensity of the fluorescent signal from the donor, reduction in the lifetime of its excited state, and/or re-emission of fluorescent light at the longer wavelengths (lower energies) characteristic of the acceptor. When the fluorescent proteins physically separate, FRET effects are diminished or eliminated. (U.S. Patent No. 5,981,200).

10 For example, a cyan fluorescent protein is excited by light at roughly 425 - 450 nm wavelength and emits light in the range of 450 - 500 nm. Yellow fluorescent protein is excited by light at roughly 500 - 525 nm and emits light at 525 - 500 nm. If these two proteins are placed in solution, the cyan and yellow fluorescence may be separately visualized. However, if these two proteins are
15 forced into close proximity with each other, the fluorescent properties will be altered by FRET. The bluish light emitted by CFP will be absorbed by YFP and re-emitted as yellow light. This means that when the proteins are stimulated with light at wavelength 450 nm, the cyan emitted light is greatly reduced and the yellow light, which is not normally stimulated at this wavelength, is greatly increased. FRET is
20 typically monitored by measuring the spectrum of emitted light in response to stimulation with light in the excitation range of the donor and calculating a ratio between the donor-emitted light and the acceptor-emitted light. When the donor:acceptor emission ratio is high, FRET is not occurring and the two fluorescent proteins are not in close proximity. When the donor: acceptor emission ratio is low,
25 FRET is occurring and the two fluorescent proteins are in close proximity. In this manner, the interaction between a first and second polypeptide may be measured.

 The occurrence of FRET also causes the fluorescence lifetime of the donor fluorescent moiety to decrease. This change in fluorescence lifetime can be measured using a technique termed fluorescence lifetime imaging technology
30 (FLIM) (Verveer et al. (2000) *Science* 290: 1567-1570; Squire et al. (1999) *J. Microsc.* 193: 36; Verveer et al. (2000) *Biophys. J.* 78: 2127). Global analysis techniques for analyzing FLIM data have been developed. These algorithms use the

understanding that the donor fluorescent moiety exists in only a limited number of states each with a distinct fluorescence lifetime. Quantitative maps of each state can be generated on a pixel-by-pixel basis.

To perform FRET-based assays, the POSH polypeptide and the interacting
 5 protein of interest are both fluorescently labeled. Suitable fluorescent labels are, in view of this specification, well known in the art. Examples are provided below, but suitable fluorescent labels not specifically discussed are also available to those of skill in the art. Fluorescent labeling may be accomplished by expressing a polypeptide as a fusion protein with a fluorescent protein, for example fluorescent
 10 proteins isolated from jellyfish, corals and other coelenterates. Exemplary fluorescent proteins include the many variants of the green fluorescent protein (GFP) of *Aequoria victoria*. Variants may be brighter, dimmer, or have different excitation and/or emission spectra. Certain variants are altered such that they no longer appear green, and may appear blue, cyan, yellow or red (termed BFP, CFP, YFP and RFP, respectively). Fluorescent proteins may be stably attached to polypeptides through a
 15 variety of covalent and noncovalent linkages, including, for example, peptide bonds (eg. expression as a fusion protein), chemical cross-linking and biotin-streptavidin coupling. For examples of fluorescent proteins, see U.S. Patents 5,625,048; 5,777,079; 6,066,476; 6,124,128; Prasher et al. (1992) *Gene*, 111:229-233; Heim et al. (1994) *Proc. Natl. Acad. Sci., USA*, 91:12501-04; Ward et al. (1982) *Photochem. Photobiol.*, 35:803-808 ; Levine et al. (1982) *Comp. Biochem. Physiol.*, 72B:77-85; Tersikh et al. (2000) *Science* 290: 1585-88.

Other exemplary fluorescent moieties well known in the art include derivatives of fluorescein, benzoxadioazole, coumarin, eosin, Lucifer Yellow,
 25 pyridyloxazole and rhodamine. These and many other exemplary fluorescent moieties may be found in the *Handbook of Fluorescent Probes and Research Chemicals* (2000, Molecular Probes, Inc.), along with methodologies for modifying polypeptides with such moieties. Exemplary proteins that fluoresce when combined with a fluorescent moiety include, yellow fluorescent protein from *Vibrio fischeri*
 30 (Baldwin et al. (1990) *Biochemistry* 29:5509-15), peridinin-chlorophyll a binding protein from the dinoflagellate *Symbiodinium sp.* (Morris et al. (1994) *Plant Molecular Biology* 24:673:77) and phycobiliproteins from marine cyanobacteria

such as *Synechococcus*, e.g., phycoerythrin and phycocyanin (Wilbanks et al. (1993) *J. Biol. Chem.* 268:1226-35). These proteins require flavins, peridinin-chlorophyll a and various phycobilins, respectively, as fluorescent co-factors.

FRET-based assays may be used in cell-based assays and in cell-free assays.

- 5 FRET-based assays are amenable to high-throughput screening methods including Fluorescence Activated Cell Sorting and fluorescent scanning of microtiter arrays.

In a further embodiment, transcript levels may be measured in cells having higher or lower levels of POSH or POSH-AP activity in order to identify genes that are regulated by POSH or POSH-APs. Promoter regions for such genes (or larger
10 portions of such genes) may be operatively linked to a reporter gene and used in a reporter gene-based assay to detect agents that enhance or diminish POSH- or POSH-AP-regulated gene expression. Transcript levels may be determined in any way known in the art, such as, for example, Northern blotting, RT-PCR, microarray, etc. Increased POSH activity may be achieved, for example, by introducing a strong
15 POSH expression vector. Decreased POSH activity may be achieved, for example, by RNAi, antisense, ribozyme, gene knockout, etc.

In general, where the screening assay is a binding assay (whether protein-protein binding, agent-protein binding, etc.), one or more of the molecules may be joined to a label, where the label can directly or indirectly provide a detectable
20 signal. Various labels include radioisotopes, fluorescers, chemiluminescers, enzymes, specific binding molecules, particles, e.g. magnetic particles, and the like. Specific binding molecules include pairs, such as biotin and streptavidin, digoxin and antidigoxin etc. For the specific binding members, the complementary member would normally be labeled with a molecule that provides for detection, in
25 accordance with known procedures.

In further embodiments, the application provides methods for identifying targets for therapeutic intervention. A polypeptide that interacts with POSH or participates in a POSH-mediated process (such as viral maturation) may be used to identify candidate therapeutics. Such targets may be identified by identifying
30 proteins that associated with POSH (POSH-APs) by, for example, immunoprecipitation with an anti-POSH antibody, in silico analysis of high-throughput binding data, two-hybrid screens, and other protein-protein interaction

assays described herein or otherwise known in the art in view of this disclosure. Agents that bind to such targets or disrupt protein-protein interactions thereof, or inhibit a biochemical activity thereof may be used in such an assay. Targets that may be identified by such approaches include: an UNC84, an MSTP028, an
 5 HERPUD1, a GTPase (eg. Rac, Rac1, Rho, Ras); an E2 enzyme, a cullin; a clathrin; AP-1; AP-2; an HSP70; an HSP90, Brca1, Bard1, Nef, PAK1, PAK2, PAK family, Vav, Cdc42, PI3K (e.g. p85 or p110), Nedd4, src (src family), Tsg101, VASP, RNB6, WASP, N-WASP, a Gag, particularly an HIV Gag (e.g. p160); and KIAA0674, Similar to Spred-2, as well as, in certain embodiments, proteins known
 10 to be associated with clathrin-coated vesicles, proteins involved in the protein sorting pathway and proteins involved in a Rac signaling pathway.

A variety of other reagents may be included in the screening assay. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc that are used to facilitate optimal protein-protein binding and/or reduce nonspecific or
 15 background interactions. Reagents that improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti- microbial agents, etc. may be used. The mixture of components are added in any order that provides for the requisite binding. Incubations are performed at any suitable temperature, typically between 4° and 40° C. Incubation periods are selected for optimum activity, but may also be
 20 optimized to facilitate rapid high-throughput screening.

In certain embodiments, a test agent may be assessed for its ability to perturb the localization of a POSH polypeptide, e.g. preventing POSH localization to the nucleus and/ or the Golgi network.

25 10. Methods and Compositions for Treatment of Viral Disorders

In a further aspect, the application provides methods and compositions for treatment of viral disorders, and particularly disorders caused by retroid viruses, RNA viruses and/or envelope viruses, including but not limited to retroviruses, rhabdoviruses, lentiviruses, and filoviruses. Preferred therapeutics of the application
 30 function by disrupting the biological activity of a POSH polypeptide or POSH complex in viral maturation. Certain therapeutics of the application function by disrupting the activity of a POSH-AP.

Exemplary therapeutics of the application include nucleic acid therapies such as for example RNAi constructs, antisense oligonucleotides, ribozyme, and DNA enzymes. Other therapeutics include polypeptides, peptidomimetics, antibodies and small molecules.

5 Antisense therapies of the application include methods of introducing antisense nucleic acids to disrupt the expression of POSH polypeptides or proteins that are necessary for POSH function, such as certain POSH-APs.

 RNAi therapies include methods of introducing RNAi constructs to downregulate the expression of POSH polypeptides or proteins that are necessary for
10 POSH function. Exemplary RNAi therapeutics include any one of SEQ ID Nos: 15, 16, 18, 19, 21, 22, 24 and 25.

 Therapeutic polypeptides may be generated by designing polypeptides to mimic certain protein domains important in the formation of POSH complexes, such as, for example SH3 or RING domains. For example, a polypeptide comprising a
15 POSH SH3 domain such as for example the SH3 domain as set forth in SEQ ID No: 30 will compete for binding to a POSH SH3 domain and will therefore act to disrupt binding of a partner protein. In one embodiment, a binding partner may be a Gag polypeptide. In another embodiment, a binding partner may be Rac. In a further embodiment, a polypeptide that resembles an L domain may disrupt recruitment of
20 Gag to the POSH complex.

 In view of the specification, methods for generating antibodies directed to epitopes of POSH and POSH-APs are known in the art. Antibodies may be introduced into cells by a variety of methods. One exemplary method comprises generating a nucleic acid encoding a single chain antibody that is capable of
25 disrupting a POSH complex. Such a nucleic acid may be conjugated to antibody that binds to receptors on the surface of target cells. It is contemplated that in certain embodiments, the antibody may target viral proteins that are present on the surface of infected cells, and in this way deliver the nucleic acid only to infected cells. Once bound to the target cell surface, the antibody is taken up by endocytosis,
30 and the conjugated nucleic acid is transcribed and translated to produce a single chain antibody that interacts with and disrupts the targeted POSH complex. Nucleic acids expressing the desired single chain antibody may also be introduced into cells

using a variety of more conventional techniques, such as viral transfection (eg. using an adenoviral system) or liposome-mediated transfection.

Small molecules of the application may be identified for their ability to modulate the formation of POSH complexes, as described above.

5 In view of the teachings herein, one of skill in the art will understand that the methods and compositions of the application are applicable to a wide range of viruses such as for example retroviral viruses, RNA viruses, and envelope viruses. In a preferred embodiment, the present application is applicable to retroviral viruses. In a more preferred embodiment, the present application is further applicable to
 10 retroviruses (retroviridae). In another more preferred embodiment, the present application is applicable to lentivirus, including primate lentivirus group. In a most preferred embodiment, the present application is applicable to Human Immunodeficiency virus (HIV), Human Immunodeficiency virus type-1 (HIV-1), Hepatitis B Virus (HBV) and Human T-cell Leukemia Virus (HTLV).

15 While not intended to be limiting, relevant retroviruses include: C-type retrovirus which causes lymphosarcoma in Northern Pike, the C-type retrovirus which infects mink, the caprine lentivirus which infects sheep, the Equine Infectious Anemia Virus (EIAV), the C-type retrovirus which infects pigs, the Avian Leukosis Sarcoma Virus (ALSV), the Feline Leukemia Virus (FeLV), the Feline Aids Virus, the Bovine Leukemia Virus (BLV), the Simian Leukemia Virus (SLV), the Simian Immuno-deficiency Virus (SIV), the Human T-cell Leukemia Virus type-I (HTLV-I), the Human T-cell Leukemia Virus type-II (HTLV-II), Human Immunodeficiency virus type-2 (HIV-2) and Human Immunodeficiency virus type-1 (HIV-1).
 20

The method and compositions of the present application are further
 25 applicable to RNA viruses, including ssRNA negative-strand viruses and ssRNA positive-strand viruses. The ssRNA positive-strand viruses include Hepatitis C Virus (HCV). In a preferred embodiment, the present application is applicable to mononegavirales, including filoviruses. Filoviruses further include Ebola viruses and Marburg viruses.

30 Other RNA viruses include picornaviruses such as enterovirus, poliovirus, coxsackievirus and hepatitis A virus, the caliciviruses, including Norwalk-like viruses, the rhabdoviruses, including rabies virus, the togaviruses including

alphaviruses, Semliki Forest virus, denguevirus, yellow fever virus and rubella virus, the orthomyxoviruses, including Type A, B, and C influenza viruses, the bunyaviruses, including the Rift Valley fever virus and the hantavirus, the filoviruses such as Ebola virus and Marburg virus, and the paramyxoviruses, including mumps virus and measles virus. Additional viruses that may be treated include herpes viruses.

11. Effective Dose

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining The Ld50 (The Dose Lethal To 50% Of The Population) And The Ed50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit large therapeutic induces are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the application, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (i.e., the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

12. Formulation and Use

Pharmaceutical compositions for use in accordance with the present application may be formulated in conventional manner using one or more physiologically acceptable carriers or excipients. Thus, the compounds and their physiologically acceptable salts and solvates may be formulated for administration
5 by, for example, injection, inhalation or insufflation (either through the mouth or the nose) or oral, buccal, parenteral or rectal administration.

An exemplary composition of the application comprises an RNAi mixed with a delivery system, such as a liposome system, and optionally including an acceptable excipient. In a preferred embodiment, the composition is formulated for
10 topical administration for, e.g. herpes virus infections.

For such therapy, the compounds of the application can be formulated for a variety of loads of administration, including systemic and topical or localized administration. Techniques and formulations generally may be found in Remington's Pharmaceutical Sciences, Meade Publishing Co., Easton, PA. For
15 systemic administration, injection is preferred, including intramuscular, intravenous, intraperitoneal, and subcutaneous. For injection, the compounds of the application can be formulated in liquid solutions, preferably in physiologically compatible buffers such as Hank's solution or Ringer's solution. In addition, the compounds may be formulated in solid form and redissolved or suspended immediately prior to
20 use. Lyophilized forms are also included.

For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (e.g., pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (e.g.,
25 lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (e.g., magnesium stearate, talc or silica); disintegrants (e.g., potato starch or sodium starch glycolate); or wetting agents (e.g., sodium lauryl sulphate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions,
30 or they may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending

agents (e.g., sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (e.g., lecithin or acacia); non-aqueous vehicles (e.g., ationd oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (e.g., methyl or propyl-p-hydroxybenzoates or sorbic acid). The preparations may also
5 contain buffer salts, flavoring, coloring and sweetening agents as appropriate.

Preparations for oral administration may be suitably formulated to give controlled release of the active compound. For buccal administration the compositions may take the form of tablets or lozenges formulated in conventional manner. For administration by inhalation, the compounds for use according to the
10 present application are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to
15 deliver a metered amount. Capsules and cartridges of e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

The compounds may be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection
20 may be presented in unit dosage form, e.g., in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle,
25 e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides.

In addition to the formulations described previously, the compounds may
30 also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with

suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

Systemic administration can also be by transmucosal or transdermal means.

- 5 For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration bile salts and fusidic acid derivatives. In addition, detergents may be used to facilitate permeation. Transmucosal administration may be through nasal sprays or using suppositories.
- 10 For topical administration, the oligomers of the application are formulated into ointments, salves, gels, or creams as generally known in the art. A wash solution can be used locally to treat an injury or inflammation to accelerate healing.

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient.

- 15 The pack may for example comprise metal or plastic foil; such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration.

- For therapies involving the administration of nucleic acids, the oligomers of the application can be formulated for a variety of modes of administration, including systemic and topical or localized administration. Techniques and formulations
- 20 generally may be found in Remington's Pharmaceutical Sciences, Meade Publishing Co., Easton, PA. For systemic administration, injection is preferred, including intramuscular, intravenous, intraperitoneal, intranodal, and subcutaneous for injection, the oligomers of the application can be formulated in liquid solutions, preferably in physiologically compatible buffers such as Hank's solution or Ringer's
- 25 solution. In addition, the oligomers may be formulated in solid form and redissolved or suspended immediately prior to use. Lyophilized forms are also included.

- Systemic administration can also be by transmucosal or transdermal means, or the compounds can be administered orally. For transmucosal or transdermal
- 30 administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration bile salts and fusidic acid derivatives. In

addition, detergents may be used to facilitate permeation. Transmucosal administration may be through nasal sprays or using suppositories. For oral administration, the oligomers are formulated into conventional oral administration forms such as capsules, tablets, and tonics. For topical administration, the oligomers of the application are formulated into ointments, salves, gels, or creams as generally known in the art.

EXEMPLIFICATION

The application now being generally described, it will be more readily understood by reference to the following examples, which are included merely for purposes of illustration of certain aspects and embodiments of the present application, and are not intended to limit the application.

EXAMPLES

1. Role of POSH in virus-like particle (VLP) budding

1. Objective:

Use RNAi to inhibit POSH gene expression and compare the efficiency of viral budding and GAG expression and processing in treated and untreated cells.

2. Study Plan:

HeLa SS-6 cells are transfected with mRNA-specific RNAi in order to knockdown the target proteins. Since maximal reduction of target protein by RNAi is achieved after 48 hours, cells are transfected twice – first to reduce target mRNAs, and subsequently to express the viral Gag protein. The second transfection is performed with pNLenv (plasmid that encodes HIV) and with low amounts of RNAi to maintain the knockdown of target protein during the time of gag expression and budding of VLPs. Reduction in mRNA levels due to RNAi effect is verified by RT-PCR amplification of target mRNA.

3. Methods, Materials, Solutions

a. Methods

i. Transfections according to manufacturer's protocol and as described in procedure.

ii. Protein determined by Bradford assay.

iii. SDS-PAGE in Hoeffer miniVE electrophoresis system: Transfer in Bio-Rad mini-protean II wet transfer system. Blots visualized using Typhoon system, and ImageQuant software

5 (ABbiotech)

b. Materials

Material	Manufacturer	Catalog #	Batch #
Lipofectamine 2000 (LF2000)	Life Technologies	11668-019	1112496
OptiMEM	Life Technologies	31985-047	3063119
RNAi Lamin A/C	Self	13	
RNAi TSG101 688	Self	65	
RNAi Posh 524	Self	81	
plenv11 PTAP	Self	148	
plenv11 ATAP	Self	149	
Anti-p24 polyclonal antibody	Seramun		A-0236/5-10-01
Anti-Rabbit Cy5 conjugated antibody	Jackson	144-175-115	48715
10% acrylamide Tris-Glycine SDS-PAGE gel	Life Technologies	NP0321	1081371
Nitrocellulose membrane	Schleicher & Schuell	401353	BA-83
NuPAGE 20X transfer buffer	Life Technologies	NP0006-1	224365
0.45µm filter	Schleicher & Schuell	10462100	CS1018-1

c. Solutions

Lysis Buffer	Compound	Concentration
--------------	----------	---------------

	Tris-HCl pH 7.6	50mM
	MgCl ₂	15mM
	NaCl	150mM
	Glycerol	10%
	EDTA	1mM
	EGTA	1mM
	ASB-14 (add immediately before use)	1%
6X Sample Buffer	Tris-HCl, pH=6.8	1M
	Glycerol	30%
	SDS	10%
	DTT	9.3%
	Bromophenol Blue	0.012%
TBS-T	Tris pH=7.6	20mM
	NaCl	137mM
	Tween-20	0.1%

4. Procedure

a. Schedule

Day				
1	2	3	4	5
Plate cells	Transfection I (RNAi only)	Passage cells (1:3)	Transfection II (RNAi and pNlenv) (12:00, PM)	Extract RNA for RT-PCR (post transfection)
			Extract RNA for RT-PCR (pre-transfection)	Harvest VLPs and cells

5 b. Day 1

Plate HeLa SS-6 cells in 6-well plates (35mm wells) at concentration of 5×10^5 cells/well.

c. Day 2

2 hours before transfection replace growth medium with 2 ml growth medium
5 without antibiotics.

Transfection I:

Reaction	RNAi name	TAGDA#	Reactions	RNAi [nM]	RNAi	A	B
					[20 μ M]	OPTiMEM	LF2000 mix
					μ l	(μ l)	(μ l)
1	Lamin A/C	13	2	50	12.5	500	500
2	Lamin A/C	13	1	50	6.25	250	250
3	TSG101 688	65	2	20	5	500	500
5	Posh 524	81	2	50	12.5	500	500

Transfections:

10 Prepare LF2000 mix : 250 μ l OptiMEM + 5 μ l LF2000 for each reaction. Mix by inversion, 5 times. Incubate 5 minutes at room temperature.

Prepare RNA dilution in OptiMEM (Table 1, column A). Add LF2000 mix dropwise to diluted RNA (Table 1, column B). Mix by gentle vortex. Incubate at room temperature 25 minutes, covered with aluminum foil.

15 Add 500 μ l transfection mixture to cells dropwise and mix by rocking side to side.

Incubate overnight.

d. Day 3

Split 1:3 after 24 hours. (Plate 4 wells for each reaction, except reaction 2 which is plated into 3 wells.)

20 e. Day 4

2 hours pre-transfection replace medium with DMEM growth medium without antibiotics.

Transfection II

A B C D

Reaction	RNAi name	TAGDA#	Plasmid	Reactions	RNAi			
					Plasmid for 2.4 µg	[20µM] for	10nM	OPTiMEM LF2000 mix
				(µg/µl)	(µl)	(µl)	(µl)	(µl)
1	Lamin A/C	13	PTAP	3	3.4	3.75	750	750
2	Lamin A/C	13	ATAP	3	2.5	3.75	750	750
3	TSG101 688	65	PTAP	3	3.4	3.75	750	750
5	Posh 524	81	PTAP	3	3.4	3.75	750	750

Prepare LF2000 mix: 250µl OptiMEM + 5 µl LF2000 for each reaction. Mix by inversion, 5 times. Incubate 5 minutes at room temperature.

Prepare RNA+DNA diluted in OptiMEM (Transfection II, A+B+C)

- 5 Add LF2000 mix (Transfection II, D) to diluted RNA+DNA dropwise, mix by gentle vortex, and incubate 1h while protected from light with aluminum foil. Add LF2000 and DNA+RNA to cells, 500µl/well, mix by gentle rocking and incubate overnight.

f. Day 5

- 10 Collect samples for VLP assay (approximately 24 hours post-transfection) by the following procedure (cells from one well from each sample is taken for RNA assay, by RT-PCR).

g. Cell Extracts

- 15
 - i. Pellet floating cells by centrifugation (5min, 3000rpm at 4°C), save supernatant (continue with supernatant immediately to step h), scrape remaining cells in the medium which remains in the well, add to the corresponding floating cell pellet and centrifuge for 5 minutes, 1800rpm at 4°C.
 - ii. Wash cell pellet twice with ice-cold PBS.
 - 20 iii. Resuspend cell pellet in 100µl lysis buffer and incubate 20 minutes on ice.
 - iv. Centrifuge at 14,000rpm for 15min. Transfer supernatant to a clean tube. This is the cell extract.
 - v. Prepare 10µl of cell extract samples for SDS-PAGE by adding SDS-
25 PAGE sample buffer to 1X, and boiling for 10 minutes. Remove an aliquot

of the remaining sample for protein determination to verify total initial starting material. Save remaining cell extract at -80 °C.

- h. Purification of VLPs from cell media
 - i. Filter the supernatant from step g through a 0.45m filter.
 - 5 ii. Centrifuge supernatant at 14,000rpm at 4°C for at least 2h.
 - iii. Aspirate supernatant carefully.
 - iv. Re-suspend VLP pellet in hot (100°C warmed for 10 min at least) 1X sample buffer.
 - v. Boil samples for 10 minutes, 100°C.
 - 10 i. Western Blot analysis
 - i. Run all samples from stages A and B on Tris-Glycine SDS-PAGE 10% (120V for 1.5h.).
 - ii. Transfer samples to nitrocellulose membrane (65V for 1.5h.).
 - iii. Stain membrane with ponceau S solution.
 - 15 iv. Block with 10% low fat milk in TBS-T for 1h.
 - v. Incubate with anti p24 rabbit 1:500 in TBS-T o/n.
 - vi. Wash 3 times with TBS-T for 7min each wash.
 - vii. Incubate with secondary antibody anti rabbit cy5 1:500 for 30min.
 - viii. Wash five times for 10min in TBS-T
 - 20 ix. View in Typhoon gel imaging system (Molecular Dynamics/APBitech) for fluorescence signal.

Results are shown in Figures 11-13.

2. Exemplary POSH RT-PCR primers and siRNA duplexes

25 RT-PCR primers

	Name	Position	Sequence
Sense primer	POSH=271	271	5' CTTCCTTGCCAGCATAC 3' (SEQ ID NO:12)
Anti-sense primer	POSH=926c	926C	5' CTGCCAGCATTCCTTCAG 3' (SEQ ID NO:13)

siRNA duplexes:

siRNA No: 153
siRNA Name: POSH-230

	Position in mRNA	426-446		
	Target sequence:	5' AACAGAGGCCTTGGAAACCTG 3'	SEQ ID NO: 14	
	siRNA sense strand:	5' dTdTcAGAGGCCUUGGAAACCUG 3'	SEQ ID NO: 15	
	siRNA anti-sense strand:	5'dTdTcAGGUUCCAAGGCCUCUG 3'	SEQ ID NO: 16	
5	siRNA No:	155		
	siRNA Name:	POSH-442		
	Position in mRNA	638-658		
	Target sequence:	5' AAAGAGCCTGGAGACCTTAAA 3'	SEQ ID NO: 17	
10	siRNA sense strand:	5' ddTdTAGAGCCUGGAGACCUUAAA 3'	SEQ ID NO: 18	
	siRNA anti-sense strand:	5' ddTdTUUUAAGGUCUCCAGGCUCU 3'	SEQ ID NO: 19	
	siRNA No:	157		
	siRNA Name:	POSH-U111		
15	Position in mRNA	2973-2993		
	Target sequence:	5' AAGGATTGGTATGTGACTCTG 3'	SEQ ID NO: 20	
	siRNA sense strand:	5' dTdTGGAUUGGUAUGUGACUCUG 3'	SEQ ID NO: 21	
	siRNA anti-sense strand:	5' dTdTcAGAGUCACAUACCAAUCC 3'	SEQ ID NO: 22	
20	siRNA No:	159		
	siRNA Name:	POSH-U410		
	Position in mRNA	3272-3292		
	Target sequence:	5' AAGCTGGATTATCTCCTGTTG 3'	SEQ ID NO: 23	
	siRNA sense strand:	5' ddTdTGCUGGAUUAUCCUGUUG 3'	SEQ ID NO: 24	
25	siRNA anti-sense strand:	5' ddTdTCAACAGGAGAUAAUCCAGC 3'	SEQ ID NO: 25	
	siRNA No.:	187		
	siRNA Name:	POSH-control		
	Position in mRNA:	None. Reverse to #153		
30	Target sequence:	5' AAGTCCAAAGGTTCCGGAGAC 3'	SEQ ID	
	NO: 36			

3. Effects of POSH RNAi on HIV Release: Kinetics

A1. Transfections

1. One day before transfection plate cells at a concentration of 5×10^6 cell/well in 15cm plates.
2. Two hours before transfection, replace cell media to 20ml complete DMEM without antibiotics.
3. DNA dilution: for each transfection dilute 62.5 μ l RNAi in 2.5ml OptiMEM according to the table below. RNAi stock is 20 μ M (recommended concentration: 50nM, dilution in total medium amount 1:400).
4. LF 2000 dilution: for each transfection dilute 50 μ l lipofectamine 2000 reagent in 2.5ml OptiMEM.
5. Incubate diluted RNAi and LF 2000 for 5 minutes at RT.
6. Mix the diluted RNAi with diluted LF2000 and incubated for 20-25 minutes at RT.
7. Add the mixture to the cells (drop wise) and incubate for 24 hours at 37°C in CO₂ incubator.
8. One day after RNAi transfection split cells (in complete MEM medium to 2 15cm plate and 1 well in a 6 wells plate)
9. One day after cells split perform HIV transfection according to SP 30-012-01.
10. 6 hours after HIV transfection replace medium to complete MEM medium.

* It is important to perform RT-PCR for Posh to assure complete knockdown.

A2. Total RNA purification.

1. One day after transfection, wash cells twice with sterile PBS.
2. Scrape cells in 2.3ml/200 μ l (for 15cm plate/1 well of a 6 wells plate) Tri reagent (with sterile scrapers) and freeze in -70°C (RNA purification and RT-PCR will be done by molecular biology unit) rack no. A16 – samples for RT.

Treatment	Chase time	Fraction	Labeling
-----------	------------	----------	----------

	(hours)		
Control=WT	1	Cells	A1
		VLP	A1 V
	2	Cells	A2
		VLP	A2 V
	3	Cells	A3
		VLP	A3 V
	4	Cells	A4
		VLP	A4 V
	5	Cells	A5
		VLP	A5 V
Posh+WT	1	Cells	B1
		VLP	B1 V
	2	Cells	B2
		VLP	B2 V
	3	Cells	B3
		VLP	B3 V
	4	Cells	B4
		VLP	B4 V
	5	Cells	B5
		VLP	B5 V

B. Labeling

1. Take out starvation medium, thaw and place at 37°C.
2. Scrape cells in growth medium and transfer gently into 15 ml conical tube.
3. Centrifuge to pellet cells at 1800rpm for 5 minutes at room temperature.
4. Aspirate supernatant and let tube stand for 10 sec. Remove the rest of the supernatant with a 200µl pipetman.

5. Gently add 10ml warm starvation medium and resuspend carefully with a 10ml pipette, up and down, just turning may not resolve the cell pellet).
6. Transfer cells to 10ml tube and place in the incubator for 60 minutes. Set an Eppendorf thermo mixer to 37°C.
- 5 7. Centrifuge to pellet cells at 1800rpm for 5 minutes at room temperature.
8. Aspirate supernatant and let tube stand for 10 sec. Remove the rest of the supernatant with a 200 µl pipetman.
9. Cut a 200µl tip from the end and resuspend cells (~ 1.5 10⁷ cells in 150 µl RPMI without Met, but try not to go over 250 µl if you have more cells)
- 10 gently in 150 µl starvation medium. Transfer cells to an Eppendorf tube and place in the thermo mixer. Wait 10 sec and transfer the rest of the cells from the 10 ml tube to the Eppendorf tube, if necessary add another 50 µl to splash the rest of the cells out (all specimens should have the same volume of labeling reaction!).
- 15 10. Pulse: Add 50 µl of ³⁵S-methionine (specific activity 14.2 µCi/µl), tightly cup tubes and place in thermo mixer. Set the mixing speed to the lowest possible (700 rpm) and incubate for 25 minutes.
11. Stop the pulse by adding 1ml ice-cold chase/stop medium. Shake tube very gently three times and pellet cells at 6000rpm for 6 sec.
- 20 12. Remove supernatant with a 1ml tip. Add gently 1ml ice-cold chase/stop medium to the pelleted cells and invert gently to resuspend.
13. Chase: Transfer all tubes to the thermo mixer and incubate for the required chase time (830:1,2,3,4 and 5 hours; 828: 3 hours only). At the end of total chase time, place tubes on ice, add 1ml ice-cold chase/stop and pellet cells
- 25 for 1 minute at 14,000 rpm. Remove supernatant and transfer supernatant to a second eppendorf tube. The cell pellet freeze at -80°C, until all tubes are ready.
14. Centrifuge supernatants for 2 hours at 14,000rpm, 4°C. Remove the supernatant very gently, leave 20 µl in the tube (labeled as V) and freeze at -
- 30 80°C until the end of the time course.

*** All steps are done on ice with ice-cold buffers

15. When the time course is over, remove all tubes from -80°C . Lyse VLP pellet (from step 14) and cell pellet (step 13) by adding 500 μl of lysis buffer (see solutions), resuspend well by pipeting up and down three times. Incubate on ice for 15 minutes, and spin in an eppendorf centrifuge for 15 minutes at 4°C , 14,000 rpm. Remove supernatant to a fresh tube, discard pellet.
16. Perform IP with anti-p24 sheep for all samples.

C. Immunoprecipitation

- 10 1. Preclearing: add to all samples 15 μl ImmunoPure PlusG (Pierce). Rotate for 1 hour at 4°C in a cycler, spin 5 min at 4°C , and transfer to a new tube for IP.
2. Add to all samples 20 μl of p24-protein G conjugated beads and incubate 4 hours in a cycler at 4°C .
3. Post immunoprecipitations, transfer all immunoprecipitations to a fresh tube.
- 15 4. Wash beads once with high salt buffer, once with medium salt buffer and once with low salt buffer. After each spin don't remove all solution, but leave 50 μl solution on the beads. After the last spin remove supernatant carefully with a loading tip and leave ~ 10 μl solution.
5. Add to each tube 20 μl 2x SDS sample buffer. Heat to 70°C for 10 minutes.
- 20 6. Samples were separated on 10% SDS-PAGE.
7. Fix gel in 25% ethanol and 10% acetic acid for 15 minutes.
8. Pour off the fixation solution and soak gels in Amplify solution (NAMP 100 Amersham) for 15 minutes.
9. Dry gels on warm plate ($60-80^{\circ}\text{C}$) under vacuum.
- 25 10. Expose gels to screen for 2 hours and scan.

4. Effect of siRNA against human POSH on production of infectious virus from HeLa cells

The following plan is according to the standard siRNA transfection protocol:

30 Plan:

Days

1	2	3	4	5.5
Plate	Transfect RNAi	Transfect siRNA +	Harvest virus	And start infectivity assay,
HeLa cells		Split high (1:3)	HIV-1 _{NL4-3}	take samples for Western blot

As siRNA the following was used:

- # 13 (= Lamin control)
- #153 (= human POSH Ub ligase)

5

day 3 (one days following first siRNA transfections):

- HeLa cells were harvested, split again 1:3 in fresh DMEM and seeded in T75 flasks.

day 4:

- 10 - medium was removed, fresh DMEM medium was added and cells were transfected again with siRNA in combination with one of the following HIV-1 expression plasmids:

- Env- (HIV-1_{NL4-3} (Adachi et al., 1986) with deletion in *env* gene, does not allow the expression of Env but other HIV-1 proteins), and VSV-G (CMV driven expression vector for vesicular stomatitis virus G protein). This combination allows single round infection as progeny viruses following infections with VSV-G pseudotyped viruses are free of Env glycoproteins.

- Env+ (HIV-1_{NL4-3} wild type) for multiple round of infections .

Note: Both construct Env- and Env+ contained EGFP were cloned into the *nef* open reading frame of HIV-1_{NL4-3}. This way all cells transfected or infected cells with active virus gene expression can be detected by autofluorescence (based on T. Fukunori et al. 2000, Lenardo et al., 2002)

For control one T25 flaks was transfected with GFP-N1 (a CMV expression vector for plain GFP).

25 day 5:

- the transfection efficiency was estimated by counting fluorescent cells in the GFP-N1 transfected culture using FACS analysis.

day 5.5:

- 36 hrs after second transfection, virus was harvested. Virus stocks were prepared as follows: HeLa cells were scraped and virus-containing supernatants were clarified by centrifugation (1,000 x g, 5 min) and filtered through a 0.45 µm-pore-size filter to remove residual cells and debris. Stocks were aliquoted and frozen at -80°C. For biochemical analyses virions from aliquots of supernatants were pelleted (99 min, 14,000 rpm, 4°C) and lysed (according to Ott et al., 2002). Samples of cell and virus fractions were analyzed by Western blot using anti-CA antibodies.
- For infectivity assay serial dilutions of virus stocks were prepared in RPMI medium and used to infect Jurkat cells.
- 3 days post infection the percentage of infected cells in parallel cultures was estimated by FACS analyses. Each infection experiments were set up in 3 parallel cultures in 96 well plates (based on Bolton et al., 2002).
- For control, one culture was incubated with cell culture supernatant from cell transfected with GFP-N1. No fluorescent cells were detected attesting for absence of unspecific staining of cells with GFP from the virus producer cells.

Remark: During virus production no toxic effect (other than some HIV-related cytopathic effect) was observed in human POSH transfected cultures when compared to the Lamin transfected culture.

Results, shown in Figure 19, demonstrate that knocking down POSH results in four logs reduction of HIV1 infectivity.

5. Effect of knocking down human POSH on PTAP- and PPEY-mediated viral release

On the first day of the experiment, HeLa SS6 cells were plated in a 6-well plate at 5×10^5 cells/well. The cells were transfected on the second day with LF2000 reagent and with the following RNAi's: Lamin, TSG101 and hPOSH (50nM/well). On the third day, cells were split 1:2. The cells were transfected again on the fourth day (with LF2000 reagent) with a supporting dose (10nM/well) of the same RNAi's and DNA (1µg/well): Lamin with PTAP, Lamin with ATAP, Lamin with PPEY, TSG101 with PTAP, TSG101 with PPEY, TSG101 with ATAP, hPOSH with PTAP, hPOSH with PPEY, and hPOSH with ATAP. 24 hours later

centrifuge supernatant for VLPs and extract protein from cells. Load all samples to a 12% tris-glycine gel, transfer to a nitrocellulose membrane and incubate with α p24.

Results, shown in Figure 20, demonstrate that the inhibition effect of viral release of knocking down human POSH is independent of specific HIV L-domain motif. In other words, human POSH appears to be important in both PTAP- and PPEY-mediated viral release. This suggests that inhibiting human POSH level/activity can be potentially effective against inhibiting the infectivity of any virus containing PTAP or PPEY motifs. HIV (contains PTAP) and Ebola virus (contains PPEY) are examples of such viruses.

6. Effect of RING and SH3 deletions on HIV-1 release

HeLa cells were sequentially transfected (within 48 hours) with a deletion construct of human POSH and pNLenv-1. The deletion constructs used in this experiments were: dRING (RING domain deletion), dSH₃4 (4th SH₃ domain deletion), dSH₃ 1 (1st SH₃ domain deletion) and dSH₃ 3 (3rd SH₃ domain deletion). As controls, HeLa cells were also sequentially transfected with mock and pNLenv-1, and mock and pNLenv-1ATAP respectively. 24 hours post transfection cells were subjected to pulse chase studies. Viral proteins were immunoprecipitated from cell lysates and pelleted virions, separated by SDS-PAGE and analyzed by fluorography. Positions of the two major CA products, p24 and p25, are indicated by double arrows. The rate of Pr55 processing was estimated by calculating the ratio of CA versus Pr55 detected intracellularly at different time point. The results, shown in Figure 21, show that RING deletion in hPOSH leads to complete budding inhibition compare to ATAP mutant control. SH₃4 domain deletion effect in the same way as RING deletion.

7. In-vitro assay of Human POSH self-ubiquitination

Recombinant hPOSH was incubated with ATP in the presence of E1, E2 and ubiquitin as indicated in each lane. Following incubation at 37°C for 30 minutes, reactions were terminated by addition of SDS-PAGE sample buffer. The samples were subsequently resolved on a 10% polyacrylamide gel. The separated samples were then transferred to nitrocellulose and subjected to immunoblot analysis with an

anti ubiquitin polyclonal antibody. The position of migration of molecular weight markers is indicated on the right.

Poly-Ub: Ub-hPOSHconjugates, detected as high molecular weight adducts only in reactions containing E1, E2 and ubiquitin. hPOSH-176 and hPOSH-178 are a short and a longer derivatives (respectively) of bacterially expressed hPOSH; C, control E3

preliminary steps in high-throughput screen

Objective

1. Test Ub detection with in a Ub chain as function of an E3 (HRD1) and POSH auto-Ubiquitination.
2. Test Boston Biochem reagents.

Materials

1. E1 recombinant from baculovirus
2. E2 Ubch5c from bacteria
3. Ubiquitin
4. POSH #178 (1-361) gst fusion-purified but degraded
5. POSH # 176 (1-269) gst fusion-purified but degraded
6. hsHRD1 soluble ring containing region
5. Bufferx12 (Tris 7.6 40 mM, DTT 1mM, MgCl₂ 5mM, ATP 2uM)
6. Dilution buffer (Tris 7.6 40mM, DTT 1mM, ovalbumin 1ug/ul)

protocol

	0.1ug/ul	0.5ug/ul	5ug/ul	0.4ug/ul	2.5ug/u/	0.8ug/ul	
	E1	E2	Ub	176	178	Hrd1	Bx12
-E1 (E2+176)	-----	0.5	0.5	1	-----	-----	10
-E2 (E1+176)	1	-----	0.5	1	-----	-----	9.5
-ub (E1+E2+176)	1	0.5	-----	1	-----	-----	9.5
E1+E2+176+Ub	1	0.5	0.5	1		-----	9
-E1 (E2+178)	-----	0.5	0.5	-----	1	-----	10
-E2 (E1+178)	1	-----	0.5	-----	1	-----	9.5
-ub (E1+E2+178)	1	0.5	-----	-----	1	-----	9.5
E1+E2+178+Ub	1	0.5	0.5	-----	1	-----1	9
Hrd1, E1+E2+Ub	1	0.5	0.5	-----	-----	1	8.5

*

1. Incubate for 30 minutes at 37°C.
2. Run 12% SDS PAGE gel and transfer to nitrocellulose membrane
3. Incubate with anti-Ubiquitin antibody.

5 Results, shown in Figure 22, demonstrate that human POSH has ubiquitin ligase activity.

8. Co-immunoprecipitation of hPOSH with myc-tagged activated (V12) and dominant-negative (N17) Rac1

10 HeLa cells were transfected with combinations of myc-Rac1 V12 or N17 and hPOSHdelRING-V5. 24 hours after transfection (efficiency 80% as measured by GFP) cells were collected, washed with PBS, and swollen in hypotonic lysis buffer (10mM HEPES pH=7.9, 15mM KCl, 0.1mM EDTA, 2mM MgCl₂, 1mM DTT, and protease inhibitors). Cells were lysed by 10 strokes with dounce homogenizer and

15 centrifuged 3000xg for 10 minutes to give supernatant (Fraction 1) and nuclei. Nuclei were washed with Fraction 2 buffer (0.2% NP-40, 10mM HEPES pH=7.9, 40mM KCl, 5% glycerol) to remove peripheral proteins. Nuclei were spun-down and supernatant collected (Fraction 2). Nuclear proteins were eluted in Fraction 3 buffer (20mM HEPES pH=7.9, 0.42M KCl, 25% glycerol, 0.1mM EDTA, 2mM

20 MgCl₂, 1mM DTT) by rotating 30 minutes in cold. Insoluble proteins were spun-down 14000xg and solubilized in Fraction 4 buffer (1% Fos-Choline 14, 50mM HEPES pH=7.9, 150mM NaCl, 10% glycerol, 1mM EDTA, 1.5mM MgCl₂, 2mM DTT). Half of the total extract was pre-cleared against Protein A sepharose for 1.5 hours and used for IP with 1 µg anti-myc (9E10, Roche 1-667-149) and Protein A

25 sepharose for 2 hours. Immune complexes were washed extensively, and eluted in SDS-PAGE sample buffer. Gels were run, and proteins electro-transferred to nitrocellulose for immunoblot as in figure 23. Endogenous POSH and transfected hPOSHdelRING-V5 are precipitated as a complex with Myc-Rac1 V12/N17. Results, shown in Figure 23, demonstrate that POSH co-immunoprecipitates with

30 Rac1.

9. Knock-down of hPOSH entraps HIV virus particles in intracellular vesicles.

HIV virus release was analyzed by electron microscopy following siRNA and full-length HIV plasmid (missing the envelope coding region) transfection. Mature viruses were secreted by cells transfected with HIV plasmid and non-relevant siRNA (control, lower panel). Knockdown of Tsg101 protein resulted in a budding defect, the viruses that were released had an immature phenotype (upper panel). Knockdown of hPOSH levels resulted in accumulation of viruses inside the cell in intracellular vesicles (middle panel). Results, shown in Figure 24, indicate that inhibiting hPOSH entraps HIV virus particles in intracellular vesicles. As accumulation of HIV virus particles in the cells accelerate cell death, inhibition of hPOSH therefore destroys HIV reservoir by killing cells infected with HIV.

10. POSH is localized at two sites one at the nuclei and one at the golgi (C)) . After HIV transfection recruitment of POSH to the golgi is enhanced (D)

HeLa and 293T cells were transfected with pNenv-1. 24 hours post transfection the cells were incubated with primary antibodies against POSH alongside with either anti-p24 Gag (B) or one of the organelle markers as follows: BiP- endoplasmatic reticulum (data not shown), GM130- golgi, (C,D) nucleoporin- nuclei matrix (E, F) and histone H1- nucleus (data not shown). Two different florescence labled secondary antibodies were used. Specific signals were obtained by laser-scanning confocal microscopy, and individual signals were overlaid to assess their relative positions. As shown in figure 7, POSH (A) is localized in the nucleus and partially colocalized with HIV-1 Gag following transfection outside the nucleus, presumably in the nuclear matrix (B).

11. POSH Ubiquitin Ligase Activity

Transfection is performed in 10 cm dishes of 293T cells. Cells are plated at 3.7×10^6 cells per plate 24 hours prior to transfection in DMEM high glucose with 10% fetal calf serum and no antibiotic (complete serum). The day of transfection the medium is replaced with 5 ml of fresh complete medium, 2 hours prior to the addition of transfection cocktail.

Transfections:

1. Place 3 ml of OptiMEM™ into 10 ml snap cap tube. Add 120 µl lipofectamin 2000™ (Invitrogene). Incubate for 5 minutes at room temperature.
 2. Meanwhile, place 1 ml OptiMEM™ in 3x10 ml snapcap tubes and the following DNAs (pNLenv-1, HIV-encoding plasmid; HA-Ubiquitin, a plasmid encoding ubiquitin with an HA tag; Empty plasmid, the vector used to clone POSH into it; POSH wild-type, full length POSH encoding plasmid; POSH C to A, POSH plamid with a mutation at the first cystein of the RING domain to an alanine):-
- | Tube no. | pNLenv-1 | HA-Ubiquitin | Expression plasmid |
|----------|----------|--------------|-----------------------|
| 10 A | 2 µg | 2 µg | Empty plasmid - 5 µg |
| B | 2 µg | 2 µg | POSH wild type - 5 µg |
| C | 2 µg | 2 µg | POSH C to A - 5 µg |
3. To each tube A, B and C add 1 ml from lipofectamine mix (from step 1).
 4. Add 1 ml of the transfection cocktail A (step 3) to plates number 1 and 4; transfection cocktail B to plates 2 and 5; transfection cocktail C to plates 3 and 6.
 5. Incubate for 24 hours.
 6. At the end of incubation wash cells twice with 1xPBS and scrape cells in 1 ml PBS.
 7. Spin at 500xg for 5 minutes to pellet cells. Lyse cells in 1 ml Triton solubilization buffer (50 mM HEPES-NaOH, pH 7.4, 150 mM NaCl, 5 mM MgCl₂, 0.5 mM EDTA, 0.5 mM DTT and protease inhibitor cocktail) for 10 minutes on ice and 10 minutes at 300C.
 8. Spin cells to pellet insoluble material at 8000xg for 20 minutes. Transfer supernatant to a fresh tube.
 9. Immunoprecipitate for 4 hours supernatants in tubes 1-3 with anti-HA antibody; and supernatants form tubes 4-6 with a cocktail of anti-gag antibodies.
 10. Wash beads extensively and add 1x SDS sample buffer. Separate extracts on 12.5% SDS-PAGE, transfer to nitrocellulose and immunoblot with anti-HA or anti-reverse transcriptase (RT) antibodies as specified in Figure 26.

As shown in Figure 27, cells expressing wild-type POSH show ubiquitination of the p160 Gag-Pol polypeptide. This ubiquitination is abolished by a Cys to Ala mutation in a conserved cysteine of the POSH Ring domain.

5 11. POSH Protein-protein interactions

POSH-associated proteins were identified by using a yeast two-hybrid assay.

Procedure:

Bait plasmid (GAL4-BD) was transformed into yeast strain AH109 (Clontech) and transformants were selected on defined media lacking tryptophan.

- 10 Yeast strain Y187 containing pre-transformed Hela cDNA prey(GAL4-AD) library (Clontech) was mated, according to the Clontech protocol, with a yeast strain containing the bait vector, and plated on defined media lacking tryptophan, leucine, histidine and containing 5 mM 3-amino-triazol. Colonies that grew on the selective media were tested for beta-galactosidase activity and positive clones were further
- 15 characterized. Plasmid was recovered from yeast colonies and transformed into E. coli DH5alpha strain. After ampicillin selection plasmid was prepared from bacterial colonies and transformed back into yeast strain AH109 together with bait plasmid or empty bait vector and colonies selected on defined media lacking leucine and tryptophan and then scored for growth on media lacking tryptophan, leucine,
- 20 histidine and containing 5 mM 3-amino-triazol. Only prey clones which their growth on this media was dependent on bait plasmid were scored as true hits. Prey clones were identified by amplifying cDNA insert and sequencing using vector derived primers.

Bait:

- 25 Plasmid vector: pGBK-T7 (Clontech)

Plasmid name: pPL269- pGBK-T7 GAL4 POSHdR

Protein sequence: Corresponds to aa 53-888 of POSH (RING domain deleted)

- RTLVGSGVEELPSNILLVRLLDGIKQRPWKPGPGGGSGTNCTNALRSQSSTVANCSSKDL
QSSQGGQQPRVQSWSPVVRGIPQLPCAALYNYEGKEPGDLKFSKGDIIILRRQVDENWY
30 HGEVNGIHGFFPTNFVQIIKPLPQPPPQCKALYDFEVKDKRADKDCLPFAKDDVLTIVRR
VDENWAEGLADKIGIFPISYVEFNAAKQLIEWDKPPVPGVDAGECSSAAQSSSTAPKH
SDTKKNTKKRHSFTSLTMANKSSQASQNRHSMEISPPVLISSSNPTAAARISELSGLSCS
APSQVHISTTGLIVTPPPSSPVTGFSFTFPPSDVPYQAALGTLNPPLPPPPLLAATVLAS

TPPGATAAAAAAGMGPRPMAGSTDQIAHLRPQTRPSVYVAIYPYTPRKEDELELRKGMF
 LVFERCQDQWFKGTSMTSKIGVFPNGYVAPVTRAVTNASQAKVPMSTAGQTSRGVTMVS
 PSTAGGPAQXLQGNVAGSPSVVPAVVVSAHIQTSPQAKVLLHMTGQMTVQNQARNVRT
 VAAHNQERPTAAVTPIQVQNAAGLSPASVGLSHSLASPQAPLMPGSATHTAAISISRA
 5 SAPLACAAAAPLTSPTSITSASLEAEPGRIVTVLPGLPTSPDSASSACGNSSATKPKDKS
 KKEKKGLLKLLSGASTKRKPRVSPASPTLEVELGSAELPLOGAVGPPELPPGGGHGRAGS
 CPVDGDGPVTTAVAGAALAQDAFHRKASSLDSAVPIAPPPRQACSSLGPVLNESPVVCE
 RHRVVVSYPYPQSEAELELKEGDIVFVHKKREDGWFKGTLQRNGKTGLFPGSFVENI

Library screened: Hela pretransformed library (Clontech).

- 10 Three POSH-APs were identified: UNC48B (Hs.5898 unc-84 homolog A),
 MSTP028 (Hs. 302746) and HERPUD1 (Hs.146393). Examples of each are shown
 below.

Human UNC84B mRNA sequence - var1 (public gi: 21265160)

15 GGCACGAGGGCAGTCCGGCAGGGAAGCCGCGCCTCCCTGAGCCTGACGCTGCAGAAAGAAGTGTGAT
 TGGAGTGACAGAGGAGCAGGTGCACCATCGTGAAGCAGGCCCTGCAGCGCTACAGTGAGGACCGCATC
 GGGCTGGCAGACTACGCCCTGGAGTCAGGAGGGGCCAGCGTCATCAGCACCCGATGTTCTGAGACCTACG
 AGACCAAGACGGCCCTCCTCAGCCTCTTCGGCATCCCCCTGTGGTACCACTCCAGTCACCCCGAGTCAT
 CCTCCAGCCAGATGTGCACCCAGGCAACTGCTGGGCCCTCCAGGGGCCACAGGGCTTCGCCGTGGTCCGC
 CTCTCTGCCCGCATCCGCCCCACAGCCGTTACCTTAGAGCATGTGCCCAAGGCCCTTGTCACCAACAGCA
 20 CTATCTCCAGTGCCCCAAGGACTTCGCCATCTTTGGGTTTGACGAAGACCTGCAGCAGGAGGGGACACT
 CCTTGGCAAGTTCACTTACGATCAGGAGCGCGAGCCTATTCAGACGTTTCACTTTCAGGCCCTACGATG
 GCCACGTACAGGTGGTGGAGCTGCGGATCTGACTAACTGGGGCCACCCGAGTACACCTGCATCTACC
 GCTTCAGAGTGATGGGGAGCCCGCCCACTAGCCCTGCTTACTGGTGCCTGCTGCCAGCCATCTGGGAGT
 GGGTGAACAGCACCCCGCGCTTCCGCCACACGCTTGCTCGGCGCTCTGACTCTAGGAGCACAAAGAGAG
 25 GAGCCTGTGGCCCCATGCAGATGAAAAGGACGGGCAGGGTCTCCTGAGCAGCAGGTGGCTCGAGGCGGTT
 AGCAGGCTCCAGCAGCTCCCTTCTTCCCTCTGTGCCCCGTGGCGTCTGCTTCCCCTCTGGGAGTGT
 GTATATAGTAGCATATCATGGGGGACTGGGAAGTTGGGAGAGGTAGGACCTGACTGTCTTGGCTGGGG
 TCAGGGGCTGGTGCCTGGGAGCTGATGAAGCAGGTGCCAGGGCTGTGGGAGGGGCAAGCTACGGCTGGG
 CTAGGTGAGCTGCCCTCTGCCCTGGGCAAGGAAGCGAGGCCCTCTGGGAGGAGGGTGTCTAGCTCCAGAG
 30 CAGGATGGGACTTCCCGAGGCAGGAAGCACTTGATGGAGAGCTGCCAGCTCTCCTACAAGGTTAGTGCC
 CTCCACCTAGGGAAGCCTGAACCAAGGGTCCCTGAGGGCCTTCGACAAAAGTGTGTATTTGTCCCGGGG
 AGGGTAGCAGTGGGCCATGGGCTTCTTGTGCCCTAAAGGGGACTGGCTGTGTGATCTTCAAGGGGCC
 CAGGGCCCAACCTGTAGGCTTCCCTCTGCTGGGACGGTAGTTGCTTTTCTCTCTCTGATGCTAGGTT
 GGGGCCCAACCTGTCTCCCTGTCTGCTAGGGCTGCCAGTGCCCTGAGCTTGCTTTCCACATTTCTCCC
 35 AGGGTATGGAGACCTAGACCTGTCTTTGGGGCATTAGCATCTGGGGTTATAGCAAGAAGAGTGGGGAGC
 ATGGAACCTCTGGGCTCTTGTGGGACGTTACGGGTATCGGGGTGCGAGGTCTGTCTGCACCGCCCCCA
 CATCTAACCAGGCCCTGATGTAGGGGTGCTCCGCTCAGGCTGCCCTTGGGCTCTTGACGCTCTTGTTT
 AGGTAGTCGCCCTTCTGGTTTGTCTCTGTGGGCAAGTGGTGGGGCTGGGGGAAGAGGCTGGCAGAAG
 TTACCTTGATAGGGAAGGGGAGGAGGGGACTTTAGAGCCAGCAGGCCCACTGTATTATGTATATAT
 40 TTTTCAAGGTCTGTTTTCTAACTGAAAAGCTAAGGGCTTGATTCTAGCCCGTCTGTGGGGCACTGG
 GTGATACTCAGTTTCTTGTCTTGGCCGTGGAGAGGGGCTGGGGCACTGGTTCCGGCTGTGTCTGGTGG
 TCCGGCTGCAGGGAAGGGGCAAGAAGCGGGCAGGCCTTCACTGCAGCACTGAGCCTCAAATCCGCTCTG
 GAGCATGAGGCTGGATGCAGTGGTGGTGGGCGCCCCGCTCCATCCCGAGGCAGCAGGCTTTGTTTTG
 45 CGCTCTCCTGTCAAAATGCTGCACTATTGGTTCTTAAGTTTATCTCCAGATCCCTAATTTATGCCTA
 TGCAAAAAATAATGACGCCCAAGAAAAA

Human UNC84B mRNA sequence - var2 (public gi: 6538748)

50 TCCCGGGTACACTCTCTGGAGCGCGCTGGAAGCTCTTGCTGCTGAATTTCTCCTCAACTGGCAGAAGG
 AGGCCATGCGGCTGGAACGCTGAGGCTGCGGCAAGGGGCTCCTGGCCAGGGAGGTGGTGGCTGAG
 CCACGAGGACACCTGGCGCTGCTGGAGGGCTAGTGAGCGCCGTGAAGCTGCCCTGAAGGAGGATTTT
 CGCAGGAAACTGCTGCTCGCATCCAGGAAGAACTGTCTGCCCTGAGAGCAGAGCATCAGCAAGACTCAG
 AAGACCTCTTCAAGAAGATCGTCCGGGCTCCAGGAGTCCGAGGCTCGCATCCAGCAGCTGAAGTCAGA
 GTGGCAAAGCATGACCCAGGAGTCTTCCAGGAGAGCTCTGTGAAGGAGCTGAGGCGGCTGGAGGACCAG
 55 CTGGCCGGCTGCAGCAGGAGCTGGCGGCTCTGGCACTGAAGCAGAGCTCGGTGGCGGAAGAAGTGGGCC
 TGCTGCCCCAGCAGATCCAGGCGGTGCGGGACGAGTGGAATCTCAGTTCCGGGCTGGATCAGTCAGTT

5 CCTTGCCCGAGGTGGAGGGGGCCGCTGGGGCTCCTTCAGAGAGAGGAGATGCAAGCTCAGCTGCGAGAG
CTGGAGAGCAGATCCTCACCCATGTGGCAGAGATGCAGGGCAAGTCGGCCAGGGAAGCCGCGCCTCCC
TGAGCCTGACGTGCAGAAAGAAGGTGTGATTGGAGTGACAGAGGAGCAGGTGCACCACATCGTGAAGCA
GGCCCTGCAGCGCTACAGTGAGGACCGCATCGGGCTGGCAGACTACGCCCTGGAGTCAGGAGGGGCCAGC
GTATCAGCACCCGATGTTCTGAGACCTACGAGACCAAGACGGCCCTCCTCAGCCTCTTCGGCATCCCC
TGTGTTACCACTCCCAGTCACCCCGAGTCACTCTCCAGCCAGATGTGCAACCAGGCAACTGCTGGGCCCT
CCAGGGGCCACAGGGCTTCGCCGTGGTCCGCCCTCTCTGCCCGCATCCGCCCCACAGCCGTTACCTTAGAG
CATGTGCCCAAGGCCTTGTCAACCAACAGCACTATCTCCAGTGCCCCCAAGGACTTCGCCATCTTTGGGT
10 TTAGCAAGACCTGCAGCAGGAGGGGACACTCCTTGGCAAGTTCACTTACGATCAGGACGGCGAGCCTAT
TCAGACGTTTCACTTTAGGCCCTACGATGGCCACGTACCAGGTGGTGGAGCTGCGGATCCTGACTAAC
TGGGGCCACCCGAGTACACTGCATCTACCCTTCAGAGTGCATGGGGAGCCCGCCCACTAG

Human UNC84B mRNA sequence - var3 (public gi: 20561571)

15 ACTCGCGCAGCGCGGGCCTCCGCCGCCGATTGTCCGCCCGCCCTTGTCCCGCTGCGCGCTCTTC
GCGTGCCCGCGCGCCCCGGGCCCGCGCTGTGTGCGCCTGAGCGGAGCGCCGCCCGGGGATCCCCACG
CGGAAAGGGCGCGCCCCGGCGCGCCTGGCCTCGGACGCCCGCGCGGGCTAGAAGCCGCGCGCGG
AGCAGATTCTCTCAGGGGAAGAGTCCACATCCCACTCATCATGTCCGAAGAAGCCAGCGCTCACGC
GCTACTCCAGGGTGACGATGACGGCAGCAGCAGCGGAGGGAGCTCGGTGGCTGGGAGTCAGAGCAC
CCTGTTTAAAGACAGTCTCTCAGGACCTTGAAGAGGAAATCCAGCAACATGAAGCGCCTGTCCCCAGCG
20 CCACAGCTGGGCCCCCTCTCTGATGCACACCTCTACTACAGTGAAGTCGCTGGTCCACGAGTCTGGT
TCCCACCCAGGAGCTCCCTGGAGGAAGTGCATGGTGACGCCAACTGGGGTGAGGACCTGCGGGTGCGGAG
GAGGAGAGGCACGGGTGGCTCAGAGAGCAGCAGGGCCAGCGGGCTTGTGGGGCGAAGGCCACCGAGGAC
TTCCTGGGCTCTTCTCGGGCTACTCCTCTGAGGACGACTACGTGGGCTACTCGGATGTGGACCAAGCAG
GTTCCAGCTCGCGCTCCGAAGCGCGTCTCACGGGCGGGCTCCTTACTCTGGATGGTGGCCACTTCGCC
25 AGGCCGCTCTTCAGACTTCTCTACTGGTGGGCTGGCACCACCTGGTACCGCCTGACCACAGCTGCCTCC
CTCCTTGACGCTTTCGTTTAAACAGGCGCTTCTCGTCCCTGAAGACGTTCTCTGTTCTCTGCTGCCGC
TGCTCTTGCTGAAGTGCCTGACGTAAGTGGTGGTATTTCTACCCCTATGGGCTGCAGACATTCACCC
TGCTTTGGTTTCTGGTGGGCGAGCGAAGGACAGCAGGAGGCGGATGAGGGCTGGGAAGCCAGAGACTCA
TCGCCACATTTCCAGGCTGAGCAGCGTGTATGTCCCGGTACACTCTCTGGAGCGCGCTGGAAGCTC
30 TTGCTGCTGAATTTCTCCAACTGGCAGAAAGGAGGCCATGCGGCTGGAAGCTCTGGAGCTGCGGCAAG
GGCTCTGGCCAGGGAGTGGTGGTGGCTGAGCCACGAGGACACCCTGGCGCTGCTGGAGGGGTAGTG
AGCCGCGTGAAGCTGCCCTGAAGGAGGATTTCGCAAGGAACTGCTGCTCGCATCCAGGAAGAACTGT
CTGCCCTGAGAGCAGAGCATCAGCAAGACTCAGAAGACCTCTTCAAGAAGATCGTCCGGGCTCCAGGA
GTCCGAGGCTCGCATCCAGCAGCTGAAGTCAGAGTGGCAAAGCATGACCAGGAGTCCCTCCAGGAGAGC
35 TCTGTGAAGGAGCTGAGGCGCTGGAGGACCAAGCTGGCCGCGCTGCAGCAGGAGCTGCGGCTTGGCAC
TGAAGCAGAGCTCGGTGGCGGAAGAAGTGGGCTGCTGCCCCAGCAGATCCAGGCCGTGCGGGACGACGT
GGAATCTCAGTTCCCGGCTGGATCAGTCAGTTCCTTGGCCGAGGTGGAGGGGCGCGTGGGGCTCCTT
CAGAGAGAGGAGATGAAGCTCAGCTGCGAGAGCTGGAGAGCAAGATCTCAACCATGTGGCAGAGATGC
AGGGCAAGTCCGCCAGGGAAGCCGCGCCTCCCTGAGCCTGACGCTGCAGAAAGAAGGTGTGATTGGAGT
40 GACAGAGGAGCAGGTGCACCAATCGTGAAGCAGGCCCTGCAGCGCTACAGTGAAGACCGCATCGGGCTG
GCAGACTACGCCCTGGAGTCAGGAGGGGCCAGCGTCATCAGCACCCGATGTTCTGAGACCTACGAGACCA
AGACGGCCCTCCTCAGCCTCTTGGCATCCCCCTGTGGTACCCTCCAGTCAACCCGAGTCACTCTCCA
GCCAGATGTGCACCCAGGCAACTGCTGGGCTTCCAGGGGCCACAAGGCTTCGCCGTGGTCCGCTCTCT
GCCCGCATCCGCCCCACAGCGGTTACCTTAGAGCATGTGCCCAAGGCCTTGTCAACCAACAGCACTATCT
45 CCAGTGCCCCCAAGGACTTCGCCATCTTTGGGTTTGACGAAGACCTGCAGCAGGAGGGGACACTCTTGG
CAAGTTCACTTACGATCAGGACGGCGAGCCTATTACAGCGTTTCACTTTACAGGCCCTACGATGGCCAG
TACCAGGTGGTGGAGCTGCGGATCTGACTAACTGGGGCCACCCGAGTACACCTGCATCTACCGTTCA
GAGTGCATGGGGAGCCCGCCACTAGCCCTGCTTACTGGTGCCTGCTGCCAGCCATCTGGGAGTGGGTGA
ACAGCACCCCGCGCTTCCCCACACGCTTGTGCGCGCTCTGACTTCTAGGAGCACAGAGAGGAGCCT
50 GTGGCCCCATGCAGATGAAAAGGACGGGCGAGGTCTCCTGAGCAGCAGGTGGCTCGAGGCGGTTAGCAGG
CTCCAGCAGCTCCCTTCTCTCTCTCTGTGCCCGTGGCGTCTGCTTCCCATCTGGGAGTGGTATAT
ATGTAGCATATCATGGGGACTGGGAAGTTGGGAGAGGTAGGACCTGACTGGTCTTGGCTGGGTCAGGG
GCTGGTGCCTGGGAGCTGATGAAGCAGGTGCCAGGGCTGTGGGAGGGGCAAGCTACGGCCTGGGCTAGGT
GAGCTGCCTCTGCCCTGGGCAAGGAAGCGAGGCCCTCTGGGAGGAGGTGCTTAGCTCCAGAGCAGGAT
55 GGGACTTCCCCAGGCAGGAAGCACTTGATGGAGAGCTGCCAGCTCTCCTACAAGGTTAGTGCCCTCCAC
CTAGGGAAGCATGAACCAAGGGTCCCTGAGGGCTTCGACAAAAGTGTGATTGTCTCCGGGGGAGGGTA
GCAGTGGGCCATGGGCTTCTTGTGCCCTAAAGGGGACTGGCTGTGTGATCTTCTAAGGGGCCAGGGC
CAACCTGTAGGCTTCCCTCTGTGTGGGACGCTAGTTGCTTTCTCTCTCTGATGTAGGTTGGGGCC
CACCCTCTCCCTGTTCTGTCTAGGGCTGCCAGTCCCTGAGCTTGCTTTCCACATTTCCAGGGTA
60 TGGAGACCTAGACCTGTCTTTGGGGCCATTAGCATCTGGGTTATAGCAAGAAGAGTGGGAGCATGGAA
CTCCTGGGCTCTTGTGGGACGTTACAGGTATCGGGGTGCGAGGTCTGTCTGCAACGGCCCCACATCTA
ACCAGGCCCTGATGTAGGGGTGCTCGCTCAGGCTGCCCCCTTGGGCTCTTGCACTCTTGTTCAGGTAG
TCGCCCTCTGGTTGTTCTCTGTGGGCAAGTTGGTGGGGCTGGGGGAAGAGGCTGGCAGAAAGTTACCC
TGGATAGGGAAGGGGAGGAGGGGACTTTTAGAGCCAGCAGGCCCACTGTATTATGTATATATTTTCA
65 AGGTCTGTTTTCTAACTGAAAAGCTAAGGGCTGATTCTAGCCCCGTTCTGTGGGGCACTGGGTGATA

CTCAGTTTCTTGTCTCTGGCCGTGGAGAGGGGCTGGGGCACTGGTTCCGGCTGTGTCTGGTGGTCTGGC
 TGGGGGGAAGGGGCAAGAAGCGGGGAGGCCCTTCACTGCAGCACTGAGCCTCAAATCCGCTCTGGAGCAT
 GAGGCTGGATGCAGTGGTGGTGAGGCCGCCCGCTCCATCCCGAGGCAGCCAGGGTTTGTTCGCTCT
 CCTGTACAAATGCTGCATATTGGTTCTTAAGTTTATCTCCAGATCCTAATTTATGCTATGCAAA
 5 AAATAAATGACGCCCAAGA

Human UNC84B mRNA sequence - var4 (public gi: 3327149)

GCTAAGAGGAGTCCCTGTGTGGGCAGCTGGAGCCTTCAATTTCTTCAGGGGAAGAGTCCACATCCAC
 CTCATCATGTCCCGAAGAAGCCAGCGCCTCACGCGTACTCCAGGGTGACGATGACGGCAGCAGCAGCA
 10 GCGGAGGGAGCTCGGTGGCTGGGAGTCAGAGCACCTGTGTTAAAGACAGTCCCTCTCAGGACCTTGAAGAG
 GAAATCCAGCAACATGAAGCGCCTGTCCCGACGCCACAGCTGGGCCCGTCTCTGATGCACACCTCTCC
 TACTACAGTGAAGTGGTCCAGAGTCCCTGGTTCCACCCAGGAGCTCCCTGGAGGAACTGCATGGTG
 ACGCCAACCTGGGGTGAGGACCTGCGGGTGCGGAGGAGAGGACCGGGTGGCTCAGAGAGCAGCAGGGC
 CAGCGGGCTTGTGGGGCGCAAGGCCACCGAGGACTTCTGGGCTCTTCTCGGGCTACTCTCTGAGGAC
 15 GACTACGTGGGCTACTCGGATGTGGACACAGCAGAGTTCAGCTCGCGGCTCCGAGCGCCGTCTCAGCGG
 CGGGCTCCTTACTCTGGATGGTGGCCACTTCGCCAGGCCGGCTCTTCAGACTTCTTACTGGTGGGCTGG
 CACCACCTGTACCGCTGACACAGCTGCCCTCCTCTGACGTCCTCGTTTTAACCAGCGCTTCTCG
 TCCCTGAAGACGTTCTCTGGTTCCTGTCTGCCGTGCTCTTGTGACGTGCTGACGTATGGTGGTGGT
 ATTTCTACCCCTATGGGCTGCAGACATTCACCCCTGCTTTGGTTTCTGGTGGGCGAGGAGGACAGCAG
 20 GAGGCCGATGAGGGCTGGGAAGCCAGAGACTCATCGCCACATTTCCAGGCTGAGCAGCGTGTATGTCC
 CGGGTACACTCTCTGGAGCGGCTCTGGAGAGCTCTGTCTGTGAATTTTCTCCAACCTGGCAGAAGGAGG
 CCATGCGGCTGGAACGCTCTGGAGCTGCGGCAAGGGGCTCCTGGCCAGGAGGTTGGTGGTGGCCTGAGCCA
 CGAGGACACCTGGCGCTGCTGGAGGGGCTAGTGAGCCGCGTGAAGCTGCCCTGAAGGAGGATTTCCGC
 AGGGAAACTGCTGCTCGCATCCAGGAAGAACTGTCTGCCCTGAGAGCAGAGCATCAGAAAGCTCAGAAG
 25 ACCTCTTCAAGAAGATCGTCCGGGCTCCAGGAGTCCGAGGCTCGCATCCAGCAGCTGAAGTCAGAGTG
 GCAAAGCATGACCCAGGAGTCTTCCAGGAGAGCTCTGTGAAGGAGCTGAGGCGGCTGGAGGACAGCTG
 GCCGGCTGACGAGGAGCTGGCGGCTCTGGCACTGAAGCAGAGCTCGGTGGCGGAAGAAGTGGGCTGCG
 TGCCCCAGCAGATCCAGGCCGTGCGGACAGAGCTGGAATCTCAGTTCCCGGCTGGATCAGTCAGTTCTCT
 TGCCCCAGGTTGAGGGGGCCCGCTGGGGCTCCTTCAGAGAGAGGAGATGCAAGCTCAGCTGCGAGAGCTG
 30 GAGAGCAAGATCCTCACCCATGTGGCAGAGATGCAGGGCAAGTCCGGCCAGGGAAGCCGCGGCTCCTTGA
 GCCTGACGCTGCAGAAAGAAGGTGTGATTGGAGTGACAGAGGAGCAGGTGCACCAATCGTGAAGCAGG
 CCTGCAGCGCTACAGTGAAGACCGCATCGGGCTGGCAGACTACGCCCTGGAGTCAGGAGGGGCCAGCGTC
 ATCAGCACCCGATGTTCTGAGACCTACGAGACCAAGACGGCCCTCCTCAGCCTCTTCGGCATCCCCCTGT
 GGTACCACTCCAGTCAACCCGAGTCACTCTCCAGCCAGATGTGCAAGCCAGGCAACTGCTGGGCTTCCA
 35 GGGGCCACAGGCTTCGCGGTGGTCCGCTCTCTGCCGCTCCGCCCCACAGCCGTACCTTAGAGCAT
 GTGCCCAAGGCTTGTACCCCAACAGCACTATCTCCAGTGCCCCCAAGGACTTCGCCATCTTTGGGTTTG
 ACGAAGACCTGCAGCAGGAGGGGACACTCCTTGGCAAGTTCACTTACGATCAGGACGGCGAGCCTATTCA
 GACGTTTCACTTTCAGGCCCTACGATGGCCACGTACAGGTGGTGGAGCTGCGGATCTGACTTAAGTGG
 40 GGCCACCCGAGTACACCTGCATCTACCGCTTCAGAGTGATGGGGAGCCCGCCCTAGCCCTGCTTAC
 TGGTGCCTGTGCTCCAGCCATCTGGGAGTGGGTGAACAGCACCCCGCGCTTCCCCCACAGCTTGTCTCG
 CGCTCTGACTTCTAGGAGCACAAGAGAGGAGCCTGTGGCCCCATGCAGATGAAAGGAGCGGAGGGTCT
 CCTGACAGCAGGTGGCTCGAGGCGGTAGCAGGCTCCAGAGCTCCCTTCTTCTTCTTCTCTGTGCCCG
 TGGCGTCTGCTTCCCATCTGGGAGTGTGTATATATGTAGCATATCATGGGGGACTGGGAAGTTGGGAGA
 45 GGTAGGACCTGACTGGTCTTGGCTGGGGTCAAGGGCTGGTGCCTGGGAGCTGATGAAGCAGGTGCCAGGG
 CTGTGGGAGGGGCAAGCTACGGCCTGGGCTAGGTGAGCTGCCCTTGGCCCTGGGCAAGGAAGCGAGGCC
 TCTGGGAGGAGGTGCTTAGCTCCAGAGCAGGATGGGACTTCCCCAGGCAGGAAGCACTTGATGGAGAGC
 TGCCAGCTCTCTACAAGGTTAGTGCCCTCCACCTAGGGAAGCCTGAACCAAGGGTCCCTGAGGGCCT
 TCGACAAAGTGTGATTGTCCCGGGAGGGTAGCAGTGGGCCATGGGGCTTCTGTGCCCTAAAGGGG
 50 ACTGGCTGCTGTGATCTTCTAAGGGGCCAGGGCCAACTGTAGGCTTCCCTCTGCTGGGGACGGTAG
 TTGCTTTTCTCTCTCTGATGCTAGGTGGGGGCCACCTGCTCCCTGTTCCTGCTAGGGCCTGCCAGTG
 CCCCTGAGCTTGTCTTCCACATCTCCAGGGTATGGAGACCTAGACCTGTCTTGGGGCCATTAGCATC
 TGGGGTTATAGCAAGAAGAGTGGGGAGCATGGAACCTCTGGGCTCTGTGGGGACGTTCAAGGATCGGG
 GTGCGAGGTCTGTCTGCACCGGCCCCACATCTAACAGGCCCTGATGTAGGGGTGCTCCGCTCAGGCTG
 55 CCCCCTTGGGCTCTGACGCTCTTGTTCAGGTAGTCGCCCTTCTGGTTTGTCTCTGTGGGGCAGTTGGT
 GGGGGCTGGGGGAAGAGGCTGGCAGAAGTTACCTGGATAGGGAAGGGGAGGAGGGGACTTTTAGAGCC
 AGCAGGCCCACTGTATTATGTATATATTTTCAAGGTCTGTTTTTCTAAGTGAAGGCTAAGGGCTTGA
 TTCTAGCCCCGTTCTGTGGGCACTGGGTGATACTCAGTTTCTTGTTCCTGGCCGTGGAGAGGGGCTG
 GGGCACTGTTCCGGCTGTGTCTGGTGGTCCGGCTGCAGGGAAGGGGCAAGAAGGGCGGCGAGCCCTCAC
 60 TGCAGCTAGAGCTCAAATCCGCTCTGGAGCATGAGGCTGGATGAGTGGTGGTGGTGGGCGCCGCTC
 CATCCCGAGGCAGCCAGGCTTTGTTTTGCGCTCTCTGTGCAAAATGCTGCACTATTGGTTCTTAAGTTT
 TTTATCTCCAGATCCTAATTTATGCCTATGCAAAAATAAATGACGCCCAAGAGCTG

Human UNC84B Protein sequence - var1 (public gi: 21265161)

HEGKSAREAAASLSLTQKEGVIGVTEQVHHIVKQALQRYSEDRIGLADYALESGGASVISTRCSSETYE

TKTALLSLFGIPLWYHSQSPRVILQPDVHPGNCWAFQGPQGFVVRLSARIRPTAVTLEHVPKALSPNST
ISSAPKDFAI FGFDEDLQOEGTLLGKFTYDQDGEPIQTFHFQAPTMATYQVVELRILTNWGHPEYTCIYR
FRVHGEPAH

5 Human UNC84B Protein sequence - var2 (public gi: 6538749)

SRVHSLERRLEALAAEFSSNWQKEAMRLERLELRQAGPQGGGGGLSHEDTLALLEGLVSRREAALKEDF
RRETAARIEELSALRAEHQQDSEDLFKKIVRASQSEARIQQLKSEWQSMTQESFQESSVKELRRLEDQ
LAGLQOELALALKQSSVAEEVGLLPQQIQAVRDDVESQFPWISQFLARGGGGRVGLLQREEMQAQLRE
LESKILTHVAEMQKKSAREAAASLSLTQKEGVIGVTEEQVHHIVKQALQRYSEDRIGLADYALESGGAS
10 VISTRSETYETKTALLSLFGIPLWYHSQSPRVILQPDVHPGNCWAFQGPQGFVVRLSARIRPTAVTLE
HVPKALSPNSTISSAPKDFAI FGFDEDLQOEGTLLGKFTYDQDGEPIQTFHFQAPTMATYQVVELRILTN
WGHPEYTCIYRFRVHGEPAH

15 Human UNC84B Protein sequence - var3 (public gi: 14778953)

MSRRSQRLLTRYSGDDDDGSSSSGGSSVAGSQSTLFKDSPLRTLKRKSSNMKRLSPAPQLGPSSDAHTSY
SESLVHESWFPFRSSLEELHGDANWGEDLRVRRRRTGGSESSRASGLVGRKATEDFLGSSSGYSSSEDDY
VGYSVDVQDQSSSSRLRSVSRAGSLLWMVATSPGRLFRLLYWWAGTTWYRLTTAASLLDVFLTRRFSSL
KTFLWFLPLLLLLTCLTYGAWFYFYPYGLQTFHPALVSWWAAKDSRRPDEGWEARDSSPHFQAEQVRMSRV
HSLERRLEALAAEFSSNWQKEAMRLERLELRQAGPQGGGGGLSHEDTLALLEGLVSRREAALKEDFRRE
20 TAARIEELSALRAEHQQDSEDLFKKIVRASQSEARIQQLKSEWQSMTQESFQESSVKELRRLEDQLAG
LQOELALALKQSSVAEEVGLLPQQIQAVRDDVESQFPWISQFLARGGGGRVGLLQREEMQAQLRELES
KILTHVAEMQKKSAREAAASLSLTQKEGVIGVTEEQVHHIVKQALQRYSEDRIGLADYALESGGASVIS
TRCSETYETKTALLSLFGIPLWYHSQSPRVILQPDVHPGNCWAFQGPQGFVVRLSARIRPTAVTLEHVP
KALSPNSTISSAPKDFAI FGFDEDLQOEGTLLGKFTYDQDGEPIQTFHFQAPTMATYQVVELRILTNWGH
25 PEYTCIYRFRVHGEPAH

Human UNC84B Protein sequence - var4 (public gi: 22096184)

LRGVPVWAAGAFRFSSEESTSHLIMSRRSQRLLTRYSGDDDDGSSSSGGSSVAGSQSTLFKDSPLRTLKR
KSSNMKRLSPAPQLGPSSDAHTSYSESLVHESWFPFRSSLEELHGDANWGEDLRVRRRRTGGSESSRA
30 SGLVGRKATEDFLGSSSGYSSSEDDYVGYSVDVQDQSSSSRLRSVSRAGSLLWMVATSPGRLFRLLYWWAG
TTWYRLTTAASLLDVFLTRRFSSSLKTFLWFLPLLLLLTCLTYGAWFYFYPYGLQTFHPALVSWWAAKDSR
RPDEGWEARDSSPHFQAEQVRMSRVHSLERRLEALAAEFSSNWQKEAMRLERLELRQAGPQGGGGGLSH
EDTLALLEGLVSRREAALKEDFRRETAARIEELSALRAEHQQDSEDLFKKIVRASQSEARIQQLKSEW
QSMTQESFQESSVKELRRLEDQLAGLQOELALALKQSSVAEEVGLLPQQIQAVRDDVESQFPWISQFL
35 ARGGGGRVGLLQREEMQAQLRELESKILTHVAEMQKKSAREAAASLSLTQKEGVIGVTEEQVHHIVKQAL
QRYSEDRIGLADYALESGGASVISTRSETYETKTALLSLFGIPLWYHSQSPRVILQPDVHPGNCWAFQ
GPQGFVVRLSARIRPTAVTLEHVPKALSPNSTISSAPKDFAI FGFDEDLQOEGTLLGKFTYDQDGEPIQ
TFHFQAPTMATYQVVELRILTNWGHPEYTCIYRFRVHGEPAH

40

Rat UNC84B mRNA sequence (public gi: 27662735)

ATGCTGCAGGCATCAGAGAGCAGCAACCTCTCACTGAAGCCCTCACTTGTAACCTGAACCTGATGGGGT
CAACATTCCTGGCACCAATCTCGTTCTAGGGCTTGGAGCCCCAGAACTGCAGGGCATGAGACACAACAC
TAAGCCCTGGGGCACTGGTAATACCCAGTTTCGCATTACATCACCCAGTATGCAGATGAGGAGAGGCCA
45 GGACCACTAGTGGATGGTGTGTGCGTGCTAGTGTGTGCACTCATGCACATGCACACTTGCTCACCCAC
TCTCCGGCACCTCTAGCATCTATGTGTGTGCTGTGCAGTCTGGAGGAGAAGAGTCCCCACTCTGCCTCAT
CATGTCAAGACGAAGCCAGCGCTCAGCGCTACTCCAGGATGACAAAGATGGTAGCAGCAGCAGTGGT
GCAAGCTCCGTGGCAGGAGGCCAGAGCACCGTGTTTAAGACAGTCCCTCTCAGGACTTTGAAGAGGAAAT
CCAGCAACATGAAGCGCTGTCCCCAGCTCCGCGAGCTGCCCCCCCCCTCTGACTCCCAACCTCTCTACTA
50 CAGCGAGTCTGTGGTTCGGGAGTCTTACATAGGCAGCCCCCGGGCTGTGTCCCTTGCCAGAAGCGCCCTC
CTGGATGACCACTACACAGTGAGCCCTACTGGAGTGGGGAACCTTCGGGGGAGGAGGAGGAGGAGGAAACAG
GCGGCTCTGAGAGCAGCAAGCCAATGGGCTCACCATGGAGAACAAGGCCTCAGAGGACTTTTTTGGCTC
TTCTCTCAGGCTACTCTCTGAGGATGACCTTGAGGCTACACGGACTCAGACCAGCACAGTTTCAGGGTCA
GGGTTAAGAAGTGCAGCATCTCGGGCCGGCTCTTTGTCTGGACGCTGGTCACTCTCCAGGCCGGCTCT
55 TTGGTCTTCTCTACTGGTGGGTGGGACCACTGGTACCGTCTGACAAAGCTGCGCTCCCTCTCTGGATGT
CTTCGTCTTAACAGGTCCAGGCACTTCTCACCAAACCTGAAGAGTTTCTGTGGTCTCTCTGCTCTTG
CTACTCCTGACTGGTCTGACCTACGGGCTCCAGACACTGCAACCCGCTGTGGCTCTGGTGGGAGCA
AGGAGATGAGGAGGAGCGAGCGAGGTGTGGGACACCGAGGATGCCTCTCGCACTTGACGAGGAGCG
CATTCTCTCCCGGTTCACTCTCTGGAGCGGCGCTGGAAGCCCTTGCTGCTGAGTTTTCTCCAAGTGG
60 CAGAAGGAGGCCATACGGCTGGAGCGCTGGAGCTGCGGAGGGGGCTGCTGGCCATGGAGGAGGAGTA
GCCTGAGCCATGAAGATGCCCTGTCTCTCTAGAAGGGTGGTGAGCCGCGGGAAGCTGCCCTGAAGGA
GCACTTGGCAGGGACACAGTGGCTCGTCCAGGAAGAGCTGGCTACCTGAGGGCAGAGCATCCAG
GACTCGGAAGACCTATTCAAGAAGATCGTCCAGGCCTCTCAGGAGTCTGAGGCCGTGTCCAGCAGCTGA
AGACAGAATGGCGAAGCATGGCTTACCTTGCTTCCCTCCATCTGGAAATAGCATGACCCAGGAGGCTTT

CCAGGAGAGCTCTGTGAAGGAGCTGGAGCGGCTGGAAGCCAACTGGCCGGTCTGAGGAGGAGCTGGCT
GCCCTAACTCTGAAGCAGAACTCGGTGGCAGATGAAGTGGGCTGCTGCCTCAGAAGATCCAGGCTGCCA
GGGCTGATGTGAATCCAGTTCCTGACTGGATCAGTCAGTTCCTTCTTCGAGATAGGGGCGCCCGCAG
CGGGCTCCTGCAGAGAGACGAGATGCATGCTCAGCTGCAGGAGCTGGAGAACAAGATTCTTGCCAATATG
5 GCTGAAATGCAGGGCAAGTCAGCCAGGAGGCGCAGCTTCTCTGGGACAGACACTGCAGAAAGAGGCA
TAGTTGGGGTGA CAGAGGAGCAAGTGCACCGGATCGTCAAGCAGGCCCTGCAGCGCTACAGTGAGACCG
TATTGGAATGGTGGATTATGCCCTGGAATCAGGAGGAGCCAGTGTCTCAGCACCCGTTGTTCTGAGACT
TACGAGACGAAGACAGCCCTTCTCAGCCTCTTTGGCATCCCCTTGTGGTACCCTCCAGTCAACCCGAG
10 TCATTCTGCAGCCAGATGTGCACCCAGGCAACTGCTGGGCATTCCAGGGGCCAGGGCTTCGAGTGGT
CCGCTCTCTGCTCGTATCCGCCCA CAGCGTCACTTAGAACACGTGCCCAAGGCCTTGTACCCAAAC
AGCACAATCTCCAGTGTCCCAAGGACTTCGCCATCTTTGGCTTTGATGAAGACCTGCAGCAGGAAGGGA
CACTTTTGGGCA CGTTTGCTACGAC CAGGATGGGAGCCATCCAGACCTTCTACTTCCAGGCCTCTAA
GATGGCCACATACCAAGTTGTGGAGCTTCGGATCTGACCAACTGGGGCCACCTGAATACACCTGTATC
15 TACCGCTTCCGGGTGCACGAGAGCCTGCCCACTAG

Rat UNC84B Protein sequence (public gi: 27662736)

MLQASESSNLFTEALTCTNLMSGSTFLAPI LVLGLGAPELQMRHNTKPWGTGNTQFRIHITQYADEERP
GPVVDGVCVLVLCALMHMTLAHPLSGTSSI YVLLCRSGGEESPLCLIMSRRSQRLLTRYSDNDGSSSSG
ASSVAGGQSTVFKDPSPLRTLKRKSSNMKRLSPAPQLPPPSDSHTSYYESVVRRESYIGSPRAVSLARSAL
20 LDDHLHSEPYWSGDLRGRRRRTGGSSESKANGLTMENKASEDFGSSSGYSSEDDLAGYTDSDQHS
GLRSAASRAGSFVWTLVTLPGRLFGLLYVWVGTTWYRLTTAASLLDVFLVTRSRHFSNPKSFLWFLLLL
LLLTGLTYGLQLQPAVASWAAKESRRQPEVWDTRDASSHLQAEQRILSRVHSLERRLEALAAEFSSNW
QKEAIRLERLELRQGAAGHGGSSLSHEDALSLEGLVSRREAALKEDLRDRTVARIQEELATLRAEHHQ
DSEDLFKKIVQASQSEARVQQLKTEWRSMALPCFPSPGNSMTQEAQFQESSVKELERLEAQLAGLRQELA
25 ALTLKQNSVADEVGLLPQKIQAA RADVESQFPDWISQFLLRDRGARSGLLQRDEHQAQLQLENKILANM
AEMQKSAAREAAASLGQTLQKEGIVGVTEEQVHRIVKQALQRYSEDRIGMVDYALESGGASVISTRSET
YETKTALLSLFGIPLWYHSQSPRVILQPDVHPGNCWAFQGPQGFVAVRLSARIRPTAVTLEHVPKALSFN
STISSAPKDFAI FGFDEDLQQEGTLLGTFAYDQDGEPIQTFFYQASKMATYQVVELRILTNWGHPEYTCI
30 YRFRVHGEPAH

Mouse UNC84B mRNA sequence (public gi: 25070155)

GACATTGCAACCCGCTGTGGTCTCCTGGTGGGCAGCAAAGAGAGCAGGAAGCAGCCAGAGGTGTGGGAA
TCCAGAGAGCCCTCCAGCACTTCCAGGCTGAGCAGCGCTTCTCTCCGGGTTCACTCTCTGGAGCGGC
35 GTCTGGAAGCCCTTGCTGCAGACTTTCTCCAAGTGGCAGAAGGAGGCCATACGGCTGGAACCGCTGGA
GCTGCGGCAGGGGCTGCTGGCCATGGAGGAGGAGTAGCCTGAGCCATGAAGATGCCCTGTCTCTCTA
GAAGGGTTGGTGAGCCGCCGAGGCTACCTGAAGGAGGACTTGCGCAGGGACACAGTGGCTCATATCC
AGGAAGAATTGGCTACCCTGAGGGCAGAGCATCACCAAGACTCGGAAGATCTCTTCAAGAAGATCGTCCA
GGCCTCTCAGGAGTCCGAAGCCCGAGTCCAGCAGCTGAAGACAGAATGGAAAAGCATGACCCAGGAGCC
40 TTCCAGGAGAGCTCTGTGAAGGAGCTGGGACGCTGGAAGCCAGCTGGCCAGCCTGCGGCAGGAGCTGG
CTGCCCTGACTCTGAAGCAGAACTCGGTGGCAGATGAAGTGGGCCCTGCTGCCACAGAAGATCCAGGCTGC
CAGGGCTGATGTGGAATCCCAGTTCCCTGACTGGATCAGGCAGTTCCTTCTGGAGACAGGGGTGCGCGC
AGCGGGCTCCTGCGAGAGATGAGATGACCGCTCAGCTGCAGGAGCTGGAGAACAAGATCCTTACCAAGA
TGGCTGAGATGCAGGCAAGTCAGCCAGGAGGCGCAGCGTCCCCTGGGACAGATACTGCAGAAAGAAGG
45 CATAGTTGGGGTGACAGAGGAGCAGGTGCACCGGATCGTCAAGCAGGCCCTGCAGCGCTACAGTGAGGAC
AGGATTGGAATGGTGGATTACGCCCTGGAATCAGGAGGAGCCAGTGTATCAGCACCCGCTGCTCTGAGA
CTTATGAGACCAAGACGGCACTCCTCAGCCTCTTTGGCATCCCCCTGTGGTACCCTCCAGTCACTCTCG
GGTCATTCTGCAGCCAGATGTGCACCCAGGCAACTGCTGGGCCTTCCAGGGGCCAGGGCTTTGCACTG
50 GTCCGCTCTCTGCTCGAATCCGACCTACAGCGTTACCTTAGAGCACGTGCCCAAGGCCTTGTACCCCA
ACAGCACTATCTCCAGTGTCTCCAAGGACTTCGCCATCTTTGGCTTCGATGAAGACCTGCAGCAGGAAGG
GACACTTCTTGGCACGTTTGCTACGAC CAGGATGGGAGCCCATCCAGACCTTCTATTTCAGGCCTCT
AAGATGGCCACATACCAAGTTGTGGAGCTTCGGATCTGACCAACTGGGGCCACCCTGAGTACACGTGTA
TCTACCGCTTCCGGGTGCACGAGAGCCTGCCCACTAGTCTTCTGAGATCTTGGGCCTGTGTGCTAGCCAC
GTGGGAAGCTCATCTTTCCAGCATCTACCTGCTCTGAAAATAGGTGCCCAGCACCCACCGCTGCCTG
55 CACATGCTTCTTAGCGCTGACTTCCAGGAGCAAGAGTGAAGAGGCAGCAGTAGAACTCCCTGCATAAGC
CAAGCCACAGGCTCCTATGCCTCTCTTAGTCTTCCCCCTAGAGCTAGGCGCCTACCAGGTGTATATATG
TAGCATACTTTGGGGGACTGCGATCAGAAGTGAGGTGGAGGACTGGCACCGGTTAACTCCTGGTAGGTG
CCAAAGCCATGGGCATACCAAGCCAGAGCCCAGGGATAAGCTGGGGGTCTTGTGGCCTTTAGGACAGT
GCTTAGCACTGTGCAGCTGGGACTGTCCCTGGCAAGTGCCTGGTAAAGGACACCGAGGCTGCTGAGCTCT
60 TCCAGCATGGAGGTACCATGGCCGTTAGGAGCCTGGGGTTCTTAGCAAAACGGGAGGAAACAGCTCCTG
TGTGCTCTGACGACTGTGTCTGTGCCAACCTAAGAGATCCCAGTCTGGGACCACTGCTTTTTTTCC
TGATGTGGGGTGAGGGTCTATTCTGTTTCC

Mouse UNC84B Protein sequence (public gi: 20902829)

MTQEAFOESSVKELGRLEAQLASLRQELAAALTLKQNSVADEVGLLPQKIQAARADVESQFPDWIROFLLG
 DRGARSGLLQRDEMHQAQLQELENKILTKMAEMQGSAREAAASLGQILQKEGIVGVTEEQVHRIVKQALQ
 RYSEDRIGMVDYALESGGASVISTRSETYETKTALLSLFGIPLWYHSQSPRVILQPDVHPGNCWAFQGP
 QGFAVVRLSARIRPTAVTLEHVPKALSPNSTISSAPKDFAIQFDFEDLQOEGTLLGTFAVDQDGEPIQTF
 YFQASKMATYQVVELRILTNWGHPEYTCIYRFRVHGEPAH

Drosophila UNC84B Protein sequence (public gi: 7302310)

MEVPTVRSPQREAEAIKVNMAIEQNIQKALTAEEYENILNHVNSYVQQLVELKMQHSHKELAPQQIELF
 VKLMKENLKQIMYKTELSEKDLSDLAIKLKLELQSSGGWQDGAKLSQANLEETKLIKAEVHLHESHYTI
 QLDRIDFASILLERILAAPALADFDARISLRVGELEPKESSGSSDAEVQIERLNRETAFIKLALSDKQAE
 NADLHQSISNLKLGQEDLLERIQQHELSDRRFHGLLAEIENKLSALNDSQFALLNKQIKLSLVEILGPK
 QSTAGGSAGQLDDFDLQTVWRSMFVAKDYLEQQOLLELNKRNTNNIRDEIERSSILMSDIQRLKREILL
 VVEAKHNESKALKGHIREEVVRQIVKTVLATYDADKTGLVDFALESAGGQILSTRCTESYQTKSAQISV
 FGIPLWYPTNTPRVAISPVPQGEWAFQGFPGFLGKSRVNMLKLNLSLVYVTGFTLEHIPKLSLPTGRIB
 SAPRNTFVWGLEQEKDQEPVLFQDYGQFEDNGASLQYFAVQNLDIKRPYEIVELRIETNHGHPTTYTCLYRF
 RVHGKPPAT

Human MSTP028 mRNA sequence - var1 (public gi: 14042294)

CCCCGCTCCGCCCCCGCTGGCGTGAGCTGGGTGTTCTCCTCCTCTCAGTCCGGGTTGGAGACTCC
 TGCGTCTCCGACTTTTCGTGGAAGAGATGTCAGGAGAAAGTGTGGTGAGCTCAGCGGTGCCAGCGGCTG
 CTACCCGCACCACTTCCCTCAAGGGCAGCAGCCCCAGCTCCAAATACGTGAAGCTGAATGTGGGTGGAGC
 CCTCTACTATACCACCATGCAGACGCTGACCAAGCAGGACACCATGCTGAAGGCCATGTTTACGCGGGCGC
 ATGGAAGTGCTCACCACAGTGAAGGCTGGATCCTCATTGACCGCTGTGGGAAGCACTTTGGTACGATAC
 TCAACTACCTTCGAGACGGGGCGGTGCCTTTACCCGAGAGCCCGCGGGAGATCGAGGAGCTGCTAGCAGA
 AGCCAAGTACTACCTAGTCCAAGGCTGGTGGGAAGAGTGCCAGGCGGCCCTACAAAACAAAGATACTTAT
 GAGCCTTTCTGCAAGGTCCCTGTGATCACCCTATCCAGGAAGAACAACAACTTATAGCGACTTCAAATA
 AGCCAGCCGTGAAGTTGCTCTACAACAGAAGTAACAACAAATACATATACCAGCAATTCTGACGACAA
 TATGTTGAAAAACATTGAAGCTGTTTGATAAGCTGTCTCTGCGCTTTAACGGAAGGGTCTGTTCAATAAG
 GATGTTATTGGGGATGAAATCTGCTGCTGGTCCCTTTTATGGTCAAGGCGCGGAAGATTGCTGAAGTCTGTT
 GTACCTCCATCGTCTATGCCACTGAGAAGAAACAGACCAAGGTGGAGTTTCCCGAAGCCCGGATTATGA
 GGAGACCTTGAAACATTTTGTCTGTATGAGGCCAGGATGCGCGGGGACCTGACAATGCCCTCTGAGGCC
 ACAGGCGGGGCGGGCGGGCGCTCCACCACTGGACGAGGACGAGGAGCGGGAGCGGATCGAGCGCGTGC
 GGAGGATCCACATCAAGCGCCCTGATGACCGGGCCACCTCCACAGTGAGCAGGCAAGAGACCGAGCCG
 CCCTCCTCTCACGCCCCCACTCCCTGCGGTGCTACACCAAGATCCTGTGACGCTGCGGGCCCTTCT
 GCTTCCCTTGGAGCCTGGAGATACTTTGTAAACAGCCAGATGATTATTTTGGTATTGCTTGACAAGGCA
 AATTGATTGTCTTGACCCAGGCGTATGACCCCTGTGCTGAACAGCTGTGTCTAAGATCTCTACTTTTC
 ATGAGAATCTGAGACTCTTTGGAGCCAGGCTTTCTCGGTTCTCAGAGGAAAGTATGAATGAGTGTGAAG
 TGTATTGTGAAACTTTTGTGCAATATTTATTTTGTGGGTGTGCGCTTCTATGTGGGCTTTTGGGT
 GACACTCCCTTAAGGGTTCAGTTTGACAATTCTGAGAGTTGTCTGCAAGTTGGAGGCCACAGAGGTATC
 TGAGCTCCCTGCTTCTATTTCAATCTCCAGCCCCAGCAGGTCCACTCTGTTCTGTGTGTTTGG
 CCGGGGCACAATCCCACTGCTTTGTGACGCTGCTTTCTGCAATGTGGCTTTGGGCTAGAGCTTGTG
 ATAATTGCAACTTGTGGCAGGGGAAATATGGCTGAATGAGCGTCTAAATCGTTGAGACCAAGTCAACTTT
 GGGTGAAGGCTTTGTTTAGGGATCAAGCCTTTTGCCACCTTGGGCTGGTCTTTGGCCTGGTGTCACTG
 GGACCCCATATGTCTGCGTAGGAGCAGAACTTTCCATGGCAGTAAGTGTCCAGCTCTGTTCTGTTCTT
 TCCCAACTCCAGCCCCGTCCAGTTGTTCTCCTGATTGACCCGACTCCACTCCAGGAAGGCCATCTGACC
 CTGTGACAGGCATAGCTCATAACTACCCCTCCCTGGGATCCGCTCCTCTTCAAGCTCCTTCCCATGA
 AGCTGGGCTAAGTTCTAAGTCATTTTGTCTAGAAATTAGTGTGGCCATACCCCTTTGTCTCCAGCC
 TGGCATCCAGGCAGGACACCTTCAACCAACAGCCCCAGGGAGCTTCCCTGCTATAAACACAGACCCCC
 TTGTCTTTGCTCTGATTTTACACAGTGTAGAGTGGCCAGCAGTGAACAGGTTGAGGATGTGCGGGTAG
 ATAGATAACTTTGGGTCTGGTTGTGTCTGTGTTTCAATGTTGTTTAAAGGATATGTGTGACTGTGGGTGG
 GGACGTGTGCTTGTGGGGCACAGGTGGCGGCCCTGTGAGCCCGGCTGGGCGCAGCGCTATGTAGGA
 CGGGTGTCTCAGTGACCTACCTCCAGGCTCCTCTGCACTGCAAGGAACAGGAGTGAGTGTGACTG
 ACAGGGGTGGTTGAGACTAGACTAGGTAGAGTAGTTACCAGGAGATGTGAATGTGCGTCAGGTGATGGAT
 GGGTTTGTCAAGGGAATCGTTACCGTTTATACCAAGGTATTAACATGGGCAGCCTTTGACACATGTAT
 TCCAAAACAGGATTTATTTTCAAACGGTTTTCACAGCTTAGACTTTGTACTTACTGCCCTGCCTGTGA
 CAGTTGTATGCCCTTCAATTTGTATCCAACAGCAAAGTCTACAATAAACTTTAAACAATCATG

Human MSTP028 mRNA sequence - var2 (public gi: 13994352)

GGAGACTCTGCGTCTCCGACTTTTCATGGAAGAGATGTCAGGAGAAAGTGTGGTGAGCTCAGCGGTGC
 CAGCGGCTGCTACCCGACCACTTCCCTCAAGGGCACAGAGCCCAAGCTCCAAATACGTGAAGCTGAATGT
 GGGTGGAGCCCTCTACTATACCACCATGCAGACGCTGACCAAGCAGGACACCATGCTGAAGGCCATGTTT
 AGCGGGCGCATGGAAGTGCTCACCGACAGTGAAGGCTGGATCCTCATTGACCGCTGTGGGAGCACTTTG
 GTACGATACTCAACTACCTTCGAGACGGGGCGGTGCCTTTACCCGAGAGCCCGCGGGAGATCGAGGAGCT

GCTAGCAGAAGCCAAAGTACTACCTAGTCCAAGGCCTGGTGAAGAGTGCCAGGCGGCCCTACAAAACAAA
 GATACTTATGAGCCTTTCTGCAAGGTCCCTGTGATCACCTCATCCAAGGAAGAACAAAACTTATAGCGA
 CTTCAAATAAGCCAGCCGTGAAGTTGCTCTACAACAGAAGTAACAACAAATACTCATATACCAGCAATTC
 5 TGACGACAATATGTTGAAAAACATTGAACTGTTTGATAAGCTGTCTCTGCGCTTTAACGGAAGGCTCCTG
 TTCATAAAGGATGTTATTGGGGATGAAATCTGCTGCTGGTCTTTTATGGTCAGGGCCGGAAGATTGCTG
 AAGTCTGTGTACCTCCATCGTCTATGCCACTGAGAAGAAACAGACCAAGGTGGAGTTTCCCAGAGCCCG
 GATTTATGAGGAGACCTTGAACATTTTGTCTGATGAGGCCAGGATGGCCGGGACCTGACAAATGCGCTC
 10 CTGGAGGCCACAGGCGGGCGGGCGGGCGCTCCACACCTGGACGAGGACGAGGAGCGGGAGCCGATCG
 AGCGCGTGGGAGGATCCACATCAAGCGCCCTGATGACCGGGCCACCTTCCAGTGAGCAGGCAAGAG
 ACCGAGCCGGCCTCTCTCACCGCCCCACTCCCTGCCGTGCTACACCCAGATCCTGTGCAGGCTGCCGG
 GCCCTTCTGCTTCCCTTGGAGCCTGGAGATACTTTTGTAAACAGCCAGATGATTATTTTGGTATTGCTT
 GACAAGGCAAAATTGATTGCTTGAACCCAGGCGTATGACCCCTGTCGTGAAACAAGCTGTGTCTAAGATCT
 CTACTTTTTCATGAGAATCTGAGACTCTTTGGAGCCAGGCTTTCTCGGTTCTCAGAGGAAAAGTATGAATG
 15 AGTGTGAAGTGATGTGAGAACTTTTGTGTGCAATATTTATTTTGTGGGTGTGCACTTCCATGTGGGC
 TTTTGGGTGACACTCCCTTAAGGGTTCAAGTTTGAACAATCTGAGAGTTGTCTGAGTTGGAGGCCAAC
 AGAGGATCTGAGCTCCCTGCTTCTTATTTCAATAATCTCCAGCCCGAGCAGGTCCACTCCTGGTCTCTG
 TGTGTTTGGCCCGGGCACAATCCCACTGCTTTGCTAGACGTGCTTTCTGCCATGTGGCTTTGGGCTAG
 AGCTTGTGTATAATTCAGCTTGTGGCAGTGGAATATGGCTGAATGAGCGTCTAAATCGTTGAGACCAG
 20 TGCAACTTTGGGTGCAAGGCTTTGTTTAGGGATCAAGCCTTTGCCACCTTGGGCTGGTCTTTGGCCTGG
 TGCTCACTGGGACCCCATATGTCTGCGTAGGAGCAGAACTTTCCATGGCAGTAAGTGTCCAGCTCTGTTT
 CTGTTCTTTCCCACTCCAGCCCCGCTCAGTTGTTCTCTGATTGACCCGACTCCACTCCAGGAAGGC
 CATCTGACCTGTGACAGGCATAGCTCATAAATACCCCTCCCTGGGATCCCGCTCCTCTCAGCCTCCT
 25 TCCCATGAAGCTGGCTAACTTTCTAAGTCAATTTGCTTAGAAATTCAGTGTGGCCCATACCTTTGTGTC
 CTCCAGCCTGGCATCCAGGCAGGACACCTCACACCAGCCCGAGGAGCTTCCCTGCTATAAACA
 CAGACCCCTTGTCTTGCCTCTGATTTTACACAGTGTAGAGTGGCCAGCAGTGAACAGGTTGAGGATG
 TGCGGTAGATAGATAAATTTGGGTCTGGTTGTGCTGCTGTTTATGTTTAAAGGATATGTTGTGAC
 TGTGGGTGGGACGTGTGCTTGTGGGGCAGAGTGGCGGCCCTGCTGGAGCCCGCTGGGCGCAGCGCC
 TATGTAGGACGGGTGTTCTCAGTGACCTACCTCCAGGCTCCTCTGCACCTGCAAAGGAACAGGAGTGAG
 30 TCGTGAAGTGAAGGGTGGTTGAGACTAGACTAGGTAGAGTAGTTACCAGGAGATGTGAATGTGCGTCAG
 GTGATGGATGGGTTTGTCAAGGGAATCGTTACCGTTTATACCAAGGTATTAAATGAGGATGCTTTGA
 CACATGTATTTCAAAAACGAGTTTATATTTTCAACGGTTTTACAGCTTAGACTTTGTACTTACTGCCC
 TGCCTGTGACAGTTGTATGCTTCATTTTGTATCCAACAGCAAGTCTACAATAAACTTTAAACAAATC
 ATGAAA

35 Human MSTP028 mRNA sequence - var3 (public gi: 25303941)
 CCGGTTTGGAGACTCTGCGTCTCCGACTTTTCAAGGAAGAGATGTGAGGAGAAAGTGTGGTGAGCTC
 AGCGGTGCCAGCGGCTGCTACCCGACCACTTCTTCAAGGGCAGAGCCCAAGCTCAAATACGTGAAG
 CTGAATGTGGGTGGAGCCCTCTACTATACCACTGACAGCGCTGACCAAGCAGGACACCATGCTGAAGG
 40 CCATGTTTCAGCGGGCGCATGGAAGTGCTCACCGACAGTGAAGGCTGGATCCTCATGACCGTGTGGGAA
 GCACCTTGGTACGATCTCAACTACCTTCGAGACGGGGCGGTGCTTTACCCGAGAGCCGCGGGAGATC
 GAGGAGCTGTAGCAGAAGCCAACTACTAGTCCAAAGGCTGGTGAAGAGTGCCAGGCGGCCCTTAC
 AAAACAAAGATACCTTATGAGCCTTTCTGCAAGGTCCCTGTGATCACTCATCAAGGAAGAACAAAACT
 TATAGCGACTTCAAATAAGCCAGCCGTGAAGTTGCTCTACAACAGAAGTAAACAATACTCATATACC
 45 AGCAATCTGACGACAATATGTTGAAAAACATTGAAGTGTGATAAGCTGTCTCTGCGCTTTAACGGAA
 GGGTCTGTTTATAAAGGATGTCATTGGGGATGAAATCTGCTGCTGGTCTTTTATGGTCAGGGCCGGAA
 GATTGCTGAAGTCTGTTGTACCTCATCTGCTATGCCACTGAGAAGAAACAGACCAAGGTGGAGTTTCCC
 GAAGCCCGGATTATGAGGAGACCTTGAACATTTGCTGTATGAGGCCAGGATGGCCGGGGACCTGACA
 50 ATGCGCTCCTGGAGGCCACAGGCGGGCGGGCGGCTCCACCACTGGACGAGGACGAGGAGCGGGA
 GCGGATCGAGCGCTGCGGAGGATCCACATCAAGCGCCCTGATGACCGGGCCACCTCCACCAAGTGAGCA
 GGCAAGAGACCGAGCCGCCCTCTCTCACCGCCCCACTCCCTGCCGTGCTACCCAGATCCTGTGCAG
 GCTGCCGGGCCCCCTTCTGCTTCCCTTGGAGCCTGGAGATACTTTGTAAACAGCCAGATGATTATTTTGG
 TATTGCTTGACAAGGCAAAATTGATTGCTTGAACCCAGGCGTATGACCCCTGTCGTGAAACAAGCTGTGTC
 55 TAAGATCTCTACTTTTTCATGAGAATCTGAGACTCTTTGGAGCCAGGCTTTCTCGGTTCTCAGAGGAAAAG
 TATGAATGAGTGTGAAGTGTATGTGAGAACTTTGTTTGAATATTTATTTTGTGGGTGTGCACTTCTT
 AGGCCACAGAGGTATCTGAGCTCCCTGCTTCTTATTTCAATAATCCTCCAGCCCGAGCAGGTCCACTCCT
 GGTTCCTGTGTGTTTGGCCCGGGCACAATCCCACTGCTTTGTAGACGTGCTTTCTGCCATGTGGCTTT
 60 GGGCTTAGAGCTTGTGATAATTGACAGCTTGTGGCAGTGGAATATGGCTGAATGAGTGTCTAAATCGTT
 GAGACAGTGCACCTTTGGGTGCAAGGCTTTGTTTAGGGATCAAGCCTTTGCCACCTTGGGCTGGTCTT
 TGGCCTGGTGTCTCACTGGGACCCCATATGTCTGCGTAGGAGCAGAACTTTCCATGGCAGTAAGTGTCCAG
 CTCTGTTTCTGTTCTTTCCCACTCCAGCCCCGCTCAGTTGTTCTCTGATTGACCCGACTCCACTCC
 65 AGGAAGGCCACTGACCCGTGACAGGCATAGCTCATAAACTACCCCTCCCTGGGATCCCGCTCCTTTC
 AGCCTCCTTCCCATGAAGCTGGGCTAACTTTCTAAGTCATTTTGTGTTAGAAATTCAGTGTGGCCATAC
 CCTTTGTCTCCAGCCTGCGATCCAGGCAGGGACACCTCACACCAGCCCGAGGGAGCTTCCCTGCT
 TATAACACAGACCCCCCTTGTCTTGGCTCTGATTTTACACAGTGTAGAGTGGCCAGCAGTGAACAGGT

TGAGGATGTGCGGGTAGATAGATAACTTTGGGTCTGGTTGTGTCTGTGTTTCATGTTTGTAAAGGATG
 TGTGACTGTGGGTGGGACGTGTGCTTGTGGGGCACAGGTGGCCCTGCTGGAGCCGGCTGGGCGCAGC
 GCCTATGTAGGACGGGTGTTCTCAGTGACCTACCTCCAGGCTCCTCTGCACCTGCAAAGGAAACAGGAGT
 5 GAGTCGTGACTGACAGGGGTGGTTGAGACTAGACTAGGTAGAGTAGTTACCAGGAGATGTGAATGTGCGT
 CAGGTGATGGATGGGTTTGTCAAGGGAATCGTTACCGTTTTATACCAAAGGTATTAAATGCGGCAGCCTT
 TGACACATGTATTCCAAAAACGAGTTTATATTTTCAAACGGTTTTTACAGCTTAGACTTTGTACTTACTG
 CCCTGCTGTGACAGTTGTATGCCCTTCATTTGTATCCAACAGCAAAAGTCTACAATAAACTTTAAAAACA
 ATCATGACTGAATGTCAAAATCGTGTATTGGGCAGATGCTTTTTAAACTGTGCTGTGAGAACTTTTATA
 10 TTAGGCCATTTGGATTTTATTAAGTGCTAAGGAAAGAGGGCTTACAAAATGTTTTCGTAAATATTTTATAC
 TGTTTAAGTGTTAAACACCAACCTGTCTTTCTTTTGGGTTGAGCTTTTTTAGAAAGTGAAGTGAATGT
 TGGCCAGGAAAAATGGAAGCCATTTGTATAAATTTTTTTTGGAGCGGAGTCTTGCTCTATTGGCCAGCG
 TGGAGTGTAGTGGCACCATCTCCACTTACCACAACCTGTGCTCCTGGGTTCAAGCGATTCTGCTGCCCTC
 AGCCTCCCGAGTAGCTGGGATTGCAGGTACCCATCAGCCCATGCCAGCTAATTTGTATTTTGTAGTAGA
 15 GATGGGGTTTCAACCATGTTGGCCAGGCTGGTCTTGAACCTCTGACCTGTGATCCGACCACCTTGGCCCTC
 CCAAAGTGTCTGGGATTACAGGTGTGAGTCAACACCTGGCTGCATAGTGTTTTAAATGTTTGTGTGAAG
 AATGAGTTTGTGGAACAATTTGATTTGTCTGTGGCCTCTATGCCCTAATGAGCTAGTGTCTTCTGGCAGCTC
 CTCTACCCAACCTTGCACCTGTAGTTTGTAGTCTTTGTCTCTCTGGAATATGAACAGGTTTATAAAACAT
 TCCATGGTGAACAATCTGTGCGCTGCATTATAGCCATGAGTGAATAGACAGCATTTGGCTGTGTTCAAGCT
 20 CTGTTATTGAGTATACAAGGAACGTATTTTCTTATGTTAGCACTAAGGGCAAAAACCAATTTTATAAT
 GTAAGCACTATCCAGGTAAAAACCTGGCCCAAGATTTGGTAAAGAGATTTCATTGCAATGTAATAACTAC
 AGTTTTTTTACAAATTGGAACAGCTTTGGTGTGTGCTAATCAAGGGTTTTTTTGTGTTGTTTGTTCAAAT
 AAGCCATCTGATTGTGGTGAAGTGGGGCCCATGTCCAAGACAATTCCTGGCATATTCTGTACCCTCCCGT
 GGGGCCATCACTGTGTGGGAGCCCATTTCCCAAGTTAAAGTGTGTCTCTGTACCTTACAACAGCGATTCA
 25 GGACCCAAGTGTGAACAACACTCAGCCCCGCCCTCTGGAGCGTGTGCTGTCTTTAGGGCTCTACCCAAAGT
 CACTGTAACAGTTAAGTGTGTCAATTAACCTTTCTGTCTCTTTGCGCCATAAAAAAATGCTCAAAGTTTTA
 GATGTAGCCACTGTATGTTGTACAAACGTTGGCGACATGTAAAAATAAAGTCATAAAATGCAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human MSTP028 mRNA sequence - var4 (public gi: 16552440)
 30 AGTCCGGGTTTGGAGACTCCTGCGTCTCCGACTTTTCATGGGCCCTGACATGGCAGGTGATATCCAGGA
 CACTGTTGGGTGCCATGGAGTTGGGGAGAGTTGGCCAGAAGAGTTGGATAACCTTGAATTGAATATTGTG
 CATGACTGTGCGGTTGCTTCTGCTGTGTGCAAGCTGCCTCCTTTACTCCAGTCCCATTTACAAAAATAA
 CGCTTGTGTTTTACAGTTATAGTTGTAGTACCCATTCAATATAGAAAACTGGAAGGCTAGACAATTC
 TTTTTCAGTTTACAGGAATAGTTCAAACAAGTTATGTGCTGTCACTGCCCTGCAGCCAAAAAGCAGCAGG
 35 AGCATACCTTTAGTCAAGCAAAAGTTGGGTTTATTTTCCTTGTGTGATTTGGGTTGGGGAAGAACTGTGGGAC
 ATCTCAGAGAAGGGCTGTGGGCTTGTGTTGGGTGATTTGAGAGACAGTTCAAGAGAAGTGGGGCTTGTCTC
 TGTGTTGGATGCTGCTGGGAAGCAGGGCTAATCTGTGATTGGGTCTCAGTGATTCTTGACTTGAAAGCA
 GGAAGAAATGGAAGGAGGCTAAACCTTCTCATTGGTAAAGCAGCAGCTGTAACCTCTATTAGCCAGGATAGG
 GGATCTTTGGTCATTTTGTATTTTGGATAATGTTTATGTTTTTGTCTGTGTCCGGACATGATGACTGAA
 40 TGGTCCGTGTTTTTGTCTTGTGCGAAGGGCACAGAGTGGCCTTGTCTGAGGGTGATGTGCTGTGAAAAAT
 GTTGATGTTCAATGGGAATGGTAGGGCCAGCCGTGGGGCTACCCAGATTCAAGCAAGATTCTGCCAC
 CCTTGCACATTTCCACCTTACAGTTTACCTGTTTCAATTCAGACATGTTTGTCTGTGATACATGTGCG
 CGGATACCAATCTCACTTTCCAGGCCCTGCGTAAATCAGCCAATGATCCATTCTTTGAGATGTACAGAG
 45 AGTCAGCCATGCTATCAGGAGATGGTAGTGGATCTTGCTCTTTTGGGCAGCACTAGTCTAGGAGGTCT
 AATTTTGCATAAATCTGGTTCCAAAAGTTTCCATGTCTGTTGTTTAGTCTCAGAAAACACCTTCTCCCTA
 CAGGAAGTGATAGGAGTGCAGCTGGAATCCCAATCAACTTCATAAAGCTTATTTTCATCTGTGATGCAGC
 TGAAAAATGACACTTAGCTAGCTATTGAGTGGTACATGGCAATAAGGAAATGTAAAGAGACCTGGGCAGT
 GCTTTAGGCTGTTTGGGTGCAGCCAGGGTGTTCATGTATACAGGTGCTAGGCAGAAAGGAAGTGCTTA
 50 TAACACAAGAGTTAGGGGCAACCTTGTGCTGCAGGGTGAAGGAGGTCAGTGTATGAGGCTTTTGG
 GGTGGGTCTTGGGACAACTAGGGGATGCATGGCCCTCTTAGGGGTGATCCAATACCCAGCTCTGACC
 AGTTGTTCCCTGTAGCCAGTGTGGCCTGTGATTTTAGGAGAAGCCAGAAGTCCAGATTTTCTGTGAG
 CTCTCCTTAGTTGACCACATTTGGAAGCAAACTTTAAATGCTGTGTATGCGTGGCCCAAGCAAAACACAT
 CTGGAGGCCAGATTGAATCCACAGGCTGAAAGCAGTCAACCAGGCCTGATGTGATGACCCTGTATCCTCT
 CCACTGGCAGGAAGAGATGTGAGGAGAAAGTGTGGTGAGCTCAGCGGTGCCAGCGGCTGCTACCCGCACC
 55 ACTTCTTCAAGGGCACGAGCCCCAGCTCCAAATACGTGAAGCTGAATGTGGGTGGAGCCCTCTACTATA
 CCACCATGCAGACGCTGACCAAGCAGGACACCATGCTGAAGGCCATGTTTCAAGGGCCATGGAAGTGCT
 CACCGACAGTGAAGGCTGGATCCTCATTTGACCGCTGTGGGAAGCACTTTGGTACGATACTCACTACCTT
 CGAGACGGGGCGGTGCTTTACCCGAGAGCCGCGGGAGATCGAGGAGCTGTAGCAGAAAGCAAGTACT
 60 ACCTAGTCAAGGCTGTGGGAAGAGTGGCAGGCGGCTCAACAGAAACAAAGATACCTTATGAGCCTTT
 CTGCAAGGTCCCTGTGATCACTCATCAAGGAAGAACAAAACTTATAGCGACTTCAAATAAGCCAGCC
 GTGAAGTTGCTCTACAAAGAAAGTAAACAAATACTCATATACCAGCAATCTGACGACAAATATGTTGA
 AAAACATTGAACGTTTGTATAAGCTGTCTGTGCGCTTTAACGGAAGGGTCTGTTCAATAAGGATGTGAT
 TGGGGTGAATCTGCTGCTGGTCCCTTTATGTTGTCAGGGCCGGAAGATTGCTGAAGTCTGTTGTGCTCC
 65 ATCGTCTATGCCACTGAGAAGAAACAGACCAAGGTGGAGTTTCCCGAAGCCCGGATTATGAGGAGACCC
 TGAACATTTGCTGTATGAGGCCAGGATGGCCGGGACCTGACAATGCGCTCCTGGAGGCCACAGGCGG

GGCGGCGGGGCGCTCCCACCACCTGGACGAGGACGAGGAGCGGGAGCGGATCGAGCGCGTGCGGAGGATC
CACATCAAGCGCCCTGATGACCGGGCCACCTCCACCAGTGAGCAGGCAAGAGACCGAGCCGCCCTCCTC
TCACGCGCCCACTCCCTGCCGTGCTACACCCAGATCCTGTGCAGGCTGCCGGGCCCCCTTCTGCTTCCCT
TGGAGCCTGGAGATACTTTGTAAACAAGCCAGATGATTATTTTGGTATTGCTTGACAAGGCAAAATTGATT
5 GTCTTGACCCAGGCGTATGACCCCTGTGCTTGAACAAGCTGTGTCTAAGATCTCTACTTTTCATGAGAAT
CTGAGACTCTTTGGAGCCAGGCTTCTCGGTCTCAGAGGAAAAGTATGAATGAGTGTGAAGTGTATGTG

Human MSTP028 mRNA sequence - var5 (public gi: 21750697)

10 GCTGGCGTGAGCTGGGTGTTTCTGCCTCTCTCAGTCCGGGTTTGGAGACTCCTGCGTCTCCGACTTTT
CATGGAAGAGATGT CAGGAGAAAGTGTGGTGAGCTCAGCGGTGCAGCGGCTGCTACCCGCACCACTTCC
TTCAAGGGCAGAGCCCCAGCTCCAAATACGTGAAGCTGAATGTGGGTGGAGCCCTCTACTATACCACCA
TGCGAGCGCTGACCAAGCAGGAACCATGTGTAAGGCCATGTCCAGCGGGCGCATGGAAGTGTCTACCCGA
CAGTGAAGAACAAGATACCTTATGAGCCTTTCTGCAAGGTCCCTGTGATCACCTCATCCAAGGAAGAACA
15 AAAACTTATAGCGACTTCAATAAAGCCAGCCGTGAAGTTGCTCTACAACAGAAGTAACAACAATACTCA
TATACCAGCGATTCTGACGACAATATGTTGAAAAACATTGAACTGTTTGATAAGCTGTCTCTGCGCTTTA
ACGGAAGGGTCTGTTTATAAAGGATGTTATTGGGGATGAAATCTGCTGCTGGTCCCTTTTATGGT CAGGG
CCGGAAGATTGTGTAAGTCTGTTGTACCTCCATCGTCTATGCCACTGAGAAGAAACAGACCAAGGTGGAG
TTTCCCGAAGCCCGGATTTATGAGGAGACCCGTGAACATTTTGTGTATGAGGCCCAGGGTGGCCGGGGAC
20 CTGACAATGCGCTCCTGGAGGCCACAGGCGGGGCGGCGGGGCGTCCCACCACCTGGACGAGACGAGGA
GCGGGAGCGGATCGAGCGCGTGCGGAGGATCCACATCAAGCGCCCTGATGACCGGGCCACCTCCACCAG
TGAGCAGGCAAGAGACCGAGCCGCCCTCCTCTCACCGCCCCACTCCCTGCCGTGCTACACCAGATCCT
GTGCAGGCTGCCGGGCCCCCTTCTGCTTCCCTTGGAGCCTGGAGATACTTTTGTAAACAAGCCAGATGATTA
TTTGTGATTGCTTGACAAGGCAAAATTGATTGTCTTGACCCAGGCGTATGACCCCTGCTGTTGAACAAGC
25 TGTGTCTAAGATCTCTACTTTTCATGAGAATCTGAGACTCTTTGGAGCCAGGCTTTCTCGGTCTCAGAG
GAAAAGTATGAATGAGTGTGAAGTGTATGTGAGAACTTTGTTTCAATATTTATTTTGTGGTGTCTGCG
CTTCTGTGTGGGCTTTTGGGTGACACTCCCTTAAGGTTTCAAGTTTGAACAATCTGAGAGTTGTCTGCG
AGTTGGAGGCCACAGAGGTATCTGAGCTCCCTGCTTCTATTTCATAATCTCCAGCCCCAGGAGTCC
ACTCCTGGTTCTGTGTGTTTGGCCCCGGGCACAATCCCACTGCTTTGCTAGACGTGCTTTCTGCCATGT
30 GGCTTTGGGCTAGAGCTTGTGATAATTGCAGCTTGTGGCAGTGGAATATGGCTGAATGAGCGTCTAA
ATCGTTGAGACCAAGTGCAACTTTGGGTGCAAGGCTTTGTTTAGGGATCAAGCCTTTTGGCCACCTTGGGCT
GGTCTTTGGCCTGGTGCTCACTGGGACCCCCATATGTCTGCGTAGGAGCAGAACTTTCCATGGCAGTAAGT
GTCCAGCTCTGTTTCTGGTTCTTTCCCAACTCCAGCCCGTCCAGTTGTTCTCCTGATTGACCCGACTC
CACTCCAGGAAGGCCATCTGACCCCTGTGACAGGCATAGCTCATAAATAACCTCCCTGGGATCCCGCTC
35 CTCTTCAGCCCTCTTCCCAATGAAGCTGGGCTAACTTTCTAAGTCAATTTTGTCTAGAAATTCAGTTGGC
CCATACCTTTGTCTCTCCAGCCTGGCATCCAGGCAGGGACACCTCACACCACAGCCCCAGGAGCTT
CCCTGCTATAAACACAGACCCCTTGTCTTTGCTCTGATTTTACACAGTGTAGAGTGGCCAGCAGTGA
ACAGGTTGAGGATGTGCGGGTAGATAGATAACTTTGGGTCTGGTTTGTGTCTGTGTTTGTGTTTAA
GGGATATGTGTGACTGTGGGTGGGGACGTGTGCTTGTGGGGCACAGGTGGCGGCCCTGCTGGAGCCTGG
40 CTGGGCGCAGCGCTATGTAGACGGGTGTTCTCAGTGACCTACCTCCAGGCTCCTCTGCACTGCAAA
GGAACAGGAGTGAGTCGTGACTGACAGGGGTGGTTGAGACTAGACTAGGTAGAGTAGTTACAGGAGATG
TGAATGTGCGTCAGGTGATGATGGGTTGTCAAGGGAATCGTTACCGTTTTATACCAAGGTATTAACA
TGGGCAGCCTTTGACACATGTATTCAAACAGAGTTTATATTTCAAACGGTTTTTACAGCTTAGACTT
45 TGTACTTACTGCCCTGCTGTGACAGTTGTATGCTTCAATTTGTATCCAACAGCAAAGTCTACAATAAA
ACTTTAAACAATCATG

Human MSTP028 Protein sequence - var1 (public gi: 13994353)

MEEMSGESVSSAVPAAATRTTSFKGTSPSSKYVKLNVGALYYTMTQTLTKQDTMLKAMFSGRMEVLTDS
SEGWILIDRCGKHFGTILNYLRDGA VPLPESRREIEELLAEAKYYLVQGLVEECQAALONKDTYEPFCKV
50 PVITSSKEEQKL IATSNKPAVKLLYNRSNNKYSYTSNSDDNMLKNIELFDKLSLRFNGRVLFIKDVI GDE
ICWVSFYGQGRKIAEVCCTSIYATEKKQTKVEFPEARIEETLNILLYEAQDGRGPDNALLEATGGAAG
RSHLDEDEERERIERVRRIRHIKRPDDRAHLHQ

Human MSTP028 Protein sequence - var2 (public gi: 14042295)

55 MSGESVSSAVPAAATRTTSFKGTSPSSKYVKLNVGALYYTMTQTLTKQDTMLKAMFSGRMEVLTDS
SEGWILIDRCGKHFGTILNYLRDGA VPLPESRREIEELLAEAKYYLVQGLVEECQAALONKDTYEPFCKV
TSSKEEQKL IATSNKPAVKLLYNRSNNKYSYTSNSDDNMLKNIELFDKLSLRFNGRVLFIKDVI GDE
WSFYGQGRKIAEVCCTSIYATEKKQTKVEFPEARIEETLNILLYEAQDGRGPDNALLEATGGAAGRSH
HLEDEDEERERIERVRRIRHIKRPDDRAHLHQ

60

Mouse MSTP028 mRNA sequence (public gi: 13905271)

CTGGGTGTTTCTGCCTCTCAGTCCGGATTGGAGACTCCGGCGTCCCTCCGACTTTTCATGGAAGAGATG
TCAGGAGACAGCGTGGTGAGCTCAGCGGTGCCAGCGGCTGCCACCCGACCACCTTCTTCAAGGGTGCCA
GCCCCAGCTCCAAGTACGTGAAGCTGAACGTGGGGGGCGCTCTCTACTACACCACCATGCAGACGCTCAC

CAAGCAGGACACCATGCTGAAGGCCATGTTTCAGTGGACGCTGGAAGTGCTCACGGACAGCGAAGGCTGG
 ATCCTCATGACCGCTGTGGGAAGCACTTTGGGACAATCCTCAACTACCTGCGGAGACGGGGGTGTACCTT
 TGCTTGAGAGCCGCCGGGAGATCGAGGAGCTGTAGCCGAGGCCAAGTACTACCTGGTGCAGGGCTGTCT
 GGAAGAGTGCAGGCCGCCCTACAAAACAAAGATACTTACGAGCCCTTCTGCAAGGTACCCGTCATCACA
 5 TCCTCCAAGGAAGAACAGAGGCTTATAGCGACGTCAAAATAAGCCCGCGTGAAGCTGCTCTACAACAGGA
 GCAACAACAAGTACTCCTACACCAGCAATTCTGACGATAACATGTTGAAAAACATCGAAGTGTTCGATAA
 GCTGTCCCTGCGCTTCAACGGGCGGGTCTGTTCATTAAGGACGTCATCGGGGACGAGATCTGCTGTTGG
 TCATTTTACGGCCAGGGCCGAAAAATCGCTGAAGTCTGTTGTACCTCCATCGTCTACGCCACTGAGAAGA
 10 AGCAGACCAAGGTTGAGTTCCCTGAAGCCGCAATTATGAGGAGACCCTGAATATCTTGGCTGTACGAGGC
 CCAGGATGGCCGAGGACCTGACAAATGCCCTCCTGGAGGCCACGGGCGGGGAGCTGGACGTTCTCATCAC
 CTGACAGGAGACGAGGAGCGGGAGCGGGAGCGGATCGAGCGCGTGAGGAGGATCCATATCAAGCGCCAG
 ATGACCGGGCCACCTCCACCACTGAGCAGGCCGAGCACCTGCTTCTGCCCCTCCCTCTGCTCTGCCCC
 CGCCCCCTCAGACCCCTGTGCAGGCTTTGGGGCACCTCCCACTTCCCTTGGAGTCTGAGACACTTTTGTAA
 CAAGCCAGATGATTATTTGTTATTTCTTGACAAAGTAGATTGCTTCTGTGTTGGCCCCGGTGTGCAACC
 15 CTTCTGAACATGCTGTGCTGCGGTCCCACTCCCGTGAGAGCCTGGAGCCCAGCCTTCTGGGGCCTTGG
 AGGAAAACGTATGAATGAGTTTGGCGTGATGTGAGAACCTTGTGTCAGTATTATTTTTATGGGTGTT
 GACTACCTATTAGGGCCTCTTAGGTGACACTCCCTCAAGGACTTAGTTTGGCAGTTGGGAGGAACCTGCT
 AGTTGGCAGCCGCCACCAAGGCGTCTTGTAGCAGCCTGGCCCTGGTGTGTTCTGAGCAGGAGTCTTGT
 CTCTCCTGTGAGTCTGCTCTGTGTGCGGTTTGTACAGAGCCCTTGGGCACAGTGGAAAGGGGACGAAG
 20 TGTGCTTCTGGCCTCTCCCTGGTTCTCAGCATGGCAGGCACACTCCAGGCTGGGAGCAGACCGATTCTGT
 GCAGTTGGCTGCTCTCTCGGCTCAGTCCCTTTGATGTACCCATCCAGGAAGGGGAGCCATCGCTGCTT
 TAAGGGTCTACTCCTCTCTTGCAATGGCTTCTTTCTCATCAAGGGGGTGCCTTGGCCCTTTGGTTAGGAA
 CTGGGGTGACCGCAGACACTGGCACCAAGGTGAGGACACCACTGAGTTTGCAGCAGTTTCCCTGCAGATA
 GAGTGGGCATGCAGATACCTTTGGGTCTTGTGTGTTTCAAGGGATGTGTGTGACTGCTGGTGGGACGT
 25 GTGACTGTAGGGCGCAGGCGCTCAGCAACAAGTCAAGTACCTCGAGTCTGCTTTGCGCAGAGGAGTGT
 GGTCTGAGTAGACTAGTTTCAAGGTTAGGTGAGAGTGTGGGATGGATTGGCTGCTTGATAAAGGAAGTGT
 TGGCTATTATTTAGACCAAGGTTATTAAGTGGGCACTCTCTGACAGGTTTAAATTTCTATTAGCGAC
 GGTTTGTACAGTCTGCTTTGTACATGACAGTGTATACCTTTCTCTGTATCTGACACAAACCTATGAT
 AAATACTCAAACAATCCCAGCTGGATCGAGGAACCATGTATGGGGTAGACTTTTGTATTAAGAGCCAAGA
 30 AAGGGAGCTTACGAATCATGTTTCCCTAGCTTTTATAGTGTGTTTAAACGGCAGCCCTGCCAGTGG
 GTTCAGCTTTTCTAGGAAGTGAATGTCAGTACTGGTGTCTTCTATAGGAATGGAGAAGCCATGTATACAC
 TGTGTAATGCTCATGTGAGAATGACCTAGCGGCACAACTGACTTGCCTTGGCCCTTGGCCCTTCCGGT
 ACTGTTTGTGGCAGCTCTCTCTACCTTCTCTATCTCAAACTTGTGCTGTAGCTTTGTACTTCACT
 35 CCCAGGGATAGGAACAGACCTAGTGAAACATCCACGGTGCCTGATCTCGCTGGCAACTGAGTCCAGCTAG
 GGCCTGACCCAGCGTCAGTCTCCAAAGCTCTGCTTCCGGATTCCAAACACTGGCGTGAGGGGACGATGTC
 AGCACTCTAGATCACCATCTAGTGAGTGCCTGGTGTAGAGTGAACCTTTACTGCACACTAAGGGCTCAC
 AATTAAATAAACAGAAATAGCTTTTGTCTCATGGTAACCAAGTTCAGTGTCTGTGGGGCCACAGCCTG
 GCAGGCTGGGCCAGCTCTCTGTACCTGCTGTGGGAGATGGACCGTTTGACCTCTCTGGAGGCTGAGAC
 CCCATACCTGCTGCTCAGTGTGGTAAATCAGCCCTCCCGAGAGTGCCTGCGGCCAGAGGGCTCCACCCA
 40 CAGTCCCTGGTCACTGACGACATCACACTCTTCTGCTCTGCACTATGAAAACCTACTGGGAGTTTT
 AGAAGTCTCACTCTTGTACGAGTGTGCGGACATGTGTAATAAAGCTTAAACTCTGCTTCTAAAAA
 AAAAAAAAAAAAAA

Mouse MSTP028 Protein sequence (public gi: 13905272)

45 MEEMSGDSVVSSAVPAAATRTTSFKGASPSKSYVKLVGGALYYTMTQTLTKQDTMLKAMFSGRMEVLTD
 SEGWLIDRCGKHFGTILNLYLRDGGVPLPESRREIEELLAEAKYLVQGLLEECQAALQNKDITYEPFCKV
 PVITSSKEEQRLIATSNKPAVKLLYNRSNNKYSYTNSDDNMLKNIELFDKLSLRFNGRVLFIKDVIIGDE
 ICCWSFYQGGRKIAEVCCTSIYATEKKQTKVEFPEARIEETLNLILLYEAQDGRGPDNALLEATGGAAG
 50 RSHHLDEDEERERERIERVRRRIHIKRPDDRAHLHQ

Drosophila MSTP028 mRNA sequence (public gi: 24585830)

CTTAAATATAAATGAAATCGATGGGTGCTTTCCATATCGATATTTCTCAAATTTTGTAGCTGTCCAA
 AACAAATAAATACATACATTAAATATACCTTTAATACAATACATTGAAGTCAAATAACAATTTAAATTGC
 55 TAATCACTATTTCCGCCAATTGTCATAACGCAATGAGAACTATGACCGACTAAGCAAGTTATCTGAATGC
 AGTAGCTGAACCCCGAAAGCTAGAGAACAAAAAGTATAAATGTCCGAATCAATGTGAGGTGATCAGAAAA
 TTTTATTAAAGGACATTCTCCCAATATTTAAACCTTAATGTTGGTGGTCAATTATATACCAAT
 TGGAACTGACGAAAAACAATGACACAATGCTGAGCGCAATGTTTAGTGGTAGAATGGAAGTGTGACT
 GATTCCGAAGGATGGATTTAATCGATCGCTGTGGAATCATTTTGTATCATTTCTTAATTATTTAAGAG
 60 CATTACCGAGTTAGCTATTTCTGTGAACGGGCACTATATGCTCATCAGGAGCCAAAGCCAATTTGCCGC
 ATTCCACTTATAACTTCAAAAAGGAAGAGCAGCTTTTATTAAGTGTCTTTTAAAGCCGCTGTATTTC
 TTGTGGTTCAGCGGCAGAAATAACAAGTATTCGTACACAAGTACATCAGACGACAACTTATTAATAAATAT
 TGAATTATTGATAAGCTTTCTTTCGCTTTAACGAACGAATTTTGTATTAAAGACGTAATTGGGCCA

AGTGAATCTGCTGCTGGTCATTTTACGGACACGGAAAAAGTAGCTGAAGTGTGTGCACTTCAATAG
 TTTATGCAACTGACAGAAAGCATACTAAAGTTGAATTTCCGGAAGCTCGTATATACGAAGAAACGCTACA
 AGTCTTGCTTTATGAAAAACCGCAATGCACCAGACCAAGAGCTCATGCAGGCAACGCTTTCAGCACGAGTA
 5 GGAAGTGCAAGTGGAACCAGCATTAAATCAGTACACAAGCGATGAGGAAGAAGAACGCACTGGACTAGCGC
 GATTACGATCCAACAAGCGTAATAATCCGTCCTGATTACAAAACTCTATTACATATGATTGACATTGTT
 TTTTATAACGATTAAAAAAAATATATATATGTTTCTTTCTTAATAAAA

Drosophila MSTP028 Protein sequence (public gi: 19921650)

10 MSESMSGDHKLILKGHSSQYLKLVGGHLYYTTIGTLTKNNDTMLSAMFSGRMEVLTDSGWILIDRCGN
 HFGIILNYLRDGTVPPLPETNKEIAELLAEAKYCYITELAISERALYAHQEPKPICRIPLITSQKEEQLL
 LSVSLKPAVILVVORQNNKYSYTSTSDMNLKNIELFDKLSLRPNERILFIKDVIGPSEICCWsfYGHGK
 KVAEVCCTSIVYATDRKHTKVEFPPEARIEETLQVLLYENRNAPDQELMQATSSARVGSASGTSINQYTS
 DEEEERTGLARLRNSNKRNNPS

Human HERPUD1 mRNA sequence - var1 (public gi: 16507801)

AGAGACGTGAACGGTCGTTCAGAGATTGCGGGCGGCTGAGACGCCGCTGCCTGGCACCTAGGAGCGCA
 CGCGAGCCCCGACACCGCCGCGCCGCGCCATGGAGTCCGAGACCGAACCCGAGCCCGTCACGCTCCTGGTG
 AAGAGCCCCAACAGCGCCACCGCGACTTGGAGCTGAGTGGCGACCGCGCTGGAGTGTGGGCCACCTCA
 AGGCCACCTGAGCCGCGTCTACCCCGAGCGTCCGCGTCCAGAGGACCAGAGGTTAATTTATTCTGGGAA
 20 GCTGTTGTTGGATCACCAAATGTCTCAGGGACTTGCTTCCAAAGGAAAAACGCGCATGTTTTCATCTGGTG
 TGCAATGTGAAGAGTCTTCAAAAAATGCCAGAAATCAACGCCAAGGTGGCTGAATCCACAGAGGAGCCTG
 CTGGTCTAATCGGGGACAGTATCCTCAGGATTCTCTAAGTGATGGTTTAAAGGCAAAGGGAAGTTCTTCG
 GAACCTTTCTTCCCTGGATGGGAAAAACATCTCAAGGCATCAGTTGGGTGGTTTCCATTTAGACCGAGG
 CCGTTTCAGAACTTCCCAAATGATGGTCTCTCCTCTGACGTTGTAAATCAGGACCCCAACAATACTTAC
 25 AGGAAGGCACTGATCCTGAACTGAAGACCCCAACCACCTCCTCCAGACAGGGATGTACTAGATGGCGA
 GCAGACCAGCCCCCTCTTTATGAGCAGAGATGGCTTGCTTCAAGACTTTCTTTGCCTCTCTTCTTCCA
 GAAGGCCCCCCAGCCATCGCAAACTGATGGTGTGTGTGCTGTAGCTGTTGGAGGCTTTGACAGGAATGGA
 CTGGATCACCTGACTCCAGCTAGATTGCCTCTCCTGGACATGGCAATGATGAGTTTTTAAAAAACAGTGT
 GGATGATGATATGCTTTTGTGAGCAAGCAAAAGCAGAAACGTGAAGCCGTGATACAAATGGGTGAACAAA
 30 AAATGCCCAAGGCTTCTCATGTCTTTATTCTGAAGAGCTTTAATATATACTCTATGTAGTTTAATAAGCA
 CTGTACGTAGAAGCCCTTAGGTGTTGCATGTCTATGCTTGAGGAACCTTTTCCAAATGTGTGTGCTGCAT
 GTGTGTTTGTACATAGAAGTCATAGATGCAGAAAGTGGTCTGCTGGTACGATTTGATTCCTGTTGGAATG
 TTTAAATTAACATAAGTGTACTACTTTATATAATCAATGAAATGCTAGACATGTTTTCAGCAGGACTTTT
 CTAGGAAAGACTTATGTATAATTGCTTTTAAATGAGTGGTCTTACTTTAACTAAGGGGAACCTTTCG
 35 GAGGTGAAAACCTTTGCTGGGTTTTCTGTTCAATAAAGTTTTACTATGAATGACCCCTGAAAAA
 AA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human HERPUD1 mRNA sequence - var2 (public gi: 10441910)

40 GCTGTGTGGCCAGGCTTTTCTCAAACCTCTGAGGGCAAGCGATCCTCCACCTCAGCCTCCTGAGTAGC
 TGGGACTACAGGCATGTGCCACTAGACCTGGCTCTAAAGCATATATGACACACGAAACCAATTTATTTTT
 CATTTCACATGTTTATTTCACATATATGGTATTAGTATTCTAATGTAGTGATGCACTTAAATTTGCATT
 ATATTTCTTAGAACATCTGAACAGAGCATAGGAAATTCCTATTCTTGCCATTATCAGTTCTAACAAAAAT
 45 CTAAAAGCACTTTATCATTTTCAATTCCTGCACTGTAATTTTTTAAATGATCAAAAACAGTATCATAC
 CAAGGCTTACTTATATTGGAATACTATTTAGAAAGTTGTGGGCTGGGTTGATTTATAAACTCTGTTGG
 TCAGATGTCTGCAATGAGTAAATTTAGCACCAATTATCAGGAAGCTTTCTCACCATGACAACTTCATTGG
 AAGATTTAATGAAAGTGTAGCATACTCTAGGGAAAAAATATGAATATTTTAGCATCTATGTATTGAAAA
 TTATGTTGAATAAATGTCAGACTATTTTACATAACGTTGCTTCTGTTTAAATTTGTACGTTTCAGAGG
 50 TGGGGGTAGGAGATGTAAGCCCTTGACAGCAAAATAATTCCTTTTGCTTGATTTTCAGACAGTTGCATCA
 GCTCCTTTGTTCTGTGTTTCATGTTACCTTATTTAGGTGGCTGAATCCACAGAGGAGCCTGTGTTCTA
 ATCGGGGACAGTATCCTGAGGATTCCTCAAGTGATGGTTTAAAGCAAAGGGAAGTTCTTCGGAACCTTTC
 TTCCCTGGATGGGAAAAATCTCAAGGCCTGAAGCTGCCAGAGGCATTCGAAGGCCTGGGTCTGTTG
 TTCTCCGTTTACACACCTATGGGTGGCTTCAGCTTTCTGGTTCCAGCAGATATATGCACGACAGTACT
 55 ACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTTTGTTCACCAACCAAGTGCAACAAGAGATACC
 TGTGGTCTCTGCACCTGCTCCAGCCCTATTTCACAAACAGTTTCCAGCTGAAAAACAGCCTGCCAATCAG
 AATGCTGCTCCTCAAGTGGTGTTAATCCTGGAGCCAAATCAAAATTTGCGGATGAATGCAACAGGTGGCC
 CTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTTGGATTGGACCTATTTCAGCAGCTACATTTTC
 TGTTTTTCTCAGTATCCTCTACTTCTACTCCTCCTGAGCAGATTCTCATGGTCATGGGGGCCACCGTT
 60 GTTATGTACCTGCATCAGTTGGGTGGTTTCCATTTAGACCGAGGCGGTTTCAGAACTTCCCAAATGATG
 GTCCTCTCCTGACGTTGTAAATCAGGACCCCAACAATAACTTACAGGAAGGCACTGATGCTGAAACTGA
 AGACCCCAACCACTCCCTCCAGACAGGGATGTACTAGATGGCGAGCAGACCAGCCCTCTTTATGAGC
 ACAGCATGGCTTGTCTTCAAGACTTTCTTTGCCTCTCTCTTCCAGAAGGCCCCCGCCATCGCAAACT
 GATGTTGTTTGTGCTGTAGCTGTTGGAGGCTTTGACAGGAATGGACTGGATCACCTGACTCCAGCTAGAT

TGCCCTCTCCTGGACATGGCAATGATGAGTTTTTAAAAACAGTGTGGATGATGATATGCTTTTGTGAGCA
 AGCAAAAGCAGAAACGTGAAGCCGTGATACAAATTTGGTGAACAAAAATGCCAAGGCTTCTCATGTCTT
 TATTCTGAAGAGCTTAAATATATACTCTATGTAGTTAATAAGCACTGTACGTAGAAGGCCTTAGGTGTT
 GCATGTCTATGCTTGAGGAACTTTTCCAAATGTGTGTCTGCATGTGTGTTGTACATAGAAGTCATAG
 5 ATGCAGAAGTGGTTCTGCTGGTACGATTGATTCTGTTGGAATGTTAAATTACCTAAGTGTACTACT
 TTATATAATCAATGAAATTGCTAGACATGTTTTAGCAGGACTTTTCTAGGAAAGACTTTATGTATAATTGC
 TTTTAAAAATGCAGTGCTTTACTTTAACTAAGGGGAACTTGCGGAGGTGAAAACCTTTGCTGGGTTTT
 CTGTTCAATAAAGTTTACTATGAATGACAAAAAATAAAAAAAAAA

10 Human HERPUD1 mRNA sequence - var3 (public gi: 3005722)
 GGGCCACTCAAGGCCACCTGAGCCGCGTCTACCCCGAGCGTCCGCGTCCAGAGGACCAGAGGTTAATTT
 ATTCGGGAAGCTGTGTGGATCACCAATGTCTCAGGGACTTGCTTCCAAAGGAAAAACGGCATGTTTT
 GCATCTGGTGTGCAATGTGAAGAGTCTTCAAAAATGCCAGAAATCAACGCCAAGGTGGCTGAATCCACA
 15 GAGGAGCCCTGCTGGTTCTAATCGGGGACAGTATCCTGAGGATTCCTCAAGTGATGGTTTAAAGCAAAGGG
 AAGTCTCTCGGAACCTTTCTCCCTGGATGGGAAAAATCTCAAGGCCTGAAGCTGCCAGCAGGCATT
 CCAAGGCCTGGGTCTGGTTTCTCCGGTTACACACCTATGGGTGGCTTCAGCTTCTGTTTCCAGCAG
 ATATATGCACGACAGTACTACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTTTGTCCACCAC
 CAAGTGCACAGAGATACCTGTGGTCTCTGCACCTGCTCCAGCCCCATTTCACAACCAAGTTTCCAGCTGA
 AAACAGCCCTGCCAATCAGAATGCTGCTCCTCAAGTGGTTGTTAATCCTGGAGCCAATCAAAATTTGCGG
 20 ATGAATGCACAAGGTGGCCCTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTTGGATTGGACCT
 ATTGAGCAGCTACATTTCTGTTTTCTCAGTATCCTCTACTTCTACTCTCTCCCTGAGCAGATTCCATCAT
 GGTCTAGGGGGCCACCGTTGTTATGTACCTGCATCAGTTGGGTGGTTTCCATTAGACCGAGGCCGGTT
 CAGAACTTCCCAAATGATGGTCTCTCTGACGTTGTAAATCAGGACCCCAACAATACTTACAGGAAG
 GCCTGATCTGAACTGAAGACCCCAACCACTCCCTCCAGACAGGATGTACTAGATGGCGAGCAGAC
 25 CAGCCCTCTTTATGAGCACAGCATGGCTGTCTTCAAGACTTTCTTGCCTCTCTTCTCCAGAAGGC
 CCCCAGCCATCGCAACTGATGGTGTGTGTGCTGTAGCTGTGGAGGCTTTGACAGGAATGGACTGGAT
 CACCTGACTCCAGCTAGATTGCTCTCTGACATGGCAATGATGAGTTTTAAAAAACAGTGTGGATGA
 TGATATGCTTTTGTGAGCAAGCAAAAGCAGAAACGTGAAGCCGTGATACAAATTGGTGAACAAAAATGC
 CCAAGGCTTCTCATGTCTTTATCTGAAGAGCTTTAATATATACTCTATGTAGTTTAAATAAGCACTGTAC
 30 GTAGAAGGCCTTAGGTGTTGCATGTCTATGCTTGAGGAACTTTTCCAAATGTGTGTCTGCATGTGTGT
 TTGTACATAGAAGTCATAGATGCAGAAGTGGTCTGCTGGTACGATTGATTCTGTGTGGAATGTTTAA
 TTACACTAAGTGTAATACTTTATATAATCAATGAAATTGCTAGACATGTTTTCAGGACTTTCTAGGA
 AAGACTTATGTATAATTGCTTTTAAAAATGCAGTGCTTTACTTTAACTAAGGGGAACCTTTCGGGAGGTG
 AAAACCTTTGCTGGGTTTTCTGTTCAATAAAGTTTACTATGAATGACCTGAAAAAATAAAAAAAAAA
 35 AAAA

Human HERPUD1 mRNA sequence - var4 (public gi: 21619176)
 CCACGCGTCCGGGTCTGTGCAGAGATTGCGGGCGGCTGAGACGCCCGCTGCTGGCACTTAGGAGCGCAG
 40 CGGAGCCCCGACACCGCCGCGCCGCGCATGGAGTCCGAGACCGAACCCGAGCCCGTCACGCTCTCTGGTGA
 AGAGCCCCAACCGAGCGCCACCGCGACTGGAGTGGCGACCGCGCTGGAGTGTGGGCCACTCTCAA
 GGCCCACTGAGCCGCTCTACCCCGAGCGTCCGCGTCCAGAGGACCAGAGGTTAATTTATTCTGGGAAG
 CTGTTGTGGATCACCAATGTCTCAGGGACTTGCTTCCAAAGCAGGAAAAACGGCATGTTTTGCATCTGG
 TGTGCAATGTGAAGAGTCTTCAAAAATGCCAGAAATCAACGCCAAGGTGGCTGAATCCACAGAGGAGCC
 45 TGCTGGTTCTAATCGGGGACAGTATCCTGAGGATTCCTCAAGTGATGGTTTAAAGGCAAAGGGAAGTTCTT
 CGGAACCTTTCTCCCTGGATGGGAAAAATCTCAAGGCCTGAAGCTGCCAGCAGGCATTCCAAGGCC
 TGGGTCTCTGGTTCTCCGGTTACACACCTATGGGTGGCTTCAGCTTTCTCTGGTTCCAGCAGATATATGC
 ACGACAGTACTACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTTTGTTCACCAACCAAGTGCA
 CAAGAGATACCTGTGGTCTCTGCACCTGCTCCAGCCCTATTCAACCAAGTTTCCAGCTGAAAACCAAGC
 50 CTGCCAATCAGAATGCTGCTCTCAAGTGGTTGTTAATCCTGGAGCCAATCAAAATTTGCGGATGAATGC
 ACAAGTGGCCCTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTTGGATTGGAACCTATTCAAGCA
 GCTACATTTTCTGTTTTCTCAGTATCCTCTACTTCTACTCTCCCTGAGCAGATTCTCTCATGGTCATGG
 GGGCCACCGTTGTTATGTACCTGCATCAGTTGGGTGGTTTCCATTAGACCGAGGCCGGTTTCAAGACTT
 CCCAAATGATGGTCTCTCTGACGTTGTAAATCAGGACCCCAACAATAACTTACAGGAAGGCACTGAT
 55 CCTGAAACTGAAGACCCCAACCACTCCCTCCAGACAGGGATGTACTAGATGGCGAGCAGACCAAGCCCT
 CCTTTATGAGCACAGCATGGCTTGTCTTCAAGACTTTCTTTCCTCTCTCTTCCAGAAGGCCCCCCAGC
 TACCGCAAACTGATGGTGTGTGTCTGTAGCTGTGGAGGCTTTGACAGGAATGGACTGGATCACCTGAC
 TCCAGCTAGATTGCTCTCTGGACATGGCAATGATGAGTTTTAAAAAACAGTGTGGATGATGATATGC
 TTTTGTGAGCAAGCAAAGCAGAAACGTGAAGCCGTGATACAAATGGTGAACAAAAATGCCAAGGCTT
 60 CTCATGTCTTTATTCTGAAGAGCTTAAATATATACTCTATGTAGTTTAAATAAGCACTGTACGTAGAAGGC
 CTTAGGTGTTGCATGTCTATGCTTGAGGAACTTTTCCAAATGTGTGTCTGCATGTGTGTTTGTACATA
 GAAGTCATAGTGCAGAAGTGGTCTGCTGGTACGATTGATTCTGTTGGAATGTTTAAATTACACTAA
 GTGTACTACTTTATATAATCAATGAAATTGCTAGACATGTTTTAGCAGGACTTTTCTAGGAAAGACTTAT
 GTATAATTGCTTTTAAATGCAGTGCTTTACTTTAACTAAGGGGAACCTTTCGGGAGGTGAAAACCTTT
 GCTGGGTTTTCTGTTCAATAAAGTTTACTATGAATGACCTGAAAAAATAAAAAAAAAA

Human HERPUD1 mRNA sequence - var5 (public gi: 14249882)

AACGGTCGTTGAGAGATTGCGGGCGGCTGAGACGCGCCCTGCCTGGCACCTAGGAGCGCAGCGGAGCCC
 CGACACCGCCGCGCCGCGCCATGGAGTCCGAGACCGAACCCGAGCCCGTCACGCTCCTGGTGAAGAGCCCC
 5 AACAGCGCCACCGCGACTTGGAGCTGAGTGGCGACCGCGGCTGGAGTGTGGGCCACCTCAAGGCCACCC
 TGAGCCGCGTCTACCCGAGCGTCCGCGTCCAGAGGACCAGAGGTTAATTTATCTGGGAAGCTGTTGTT
 GGATCACCAATGTCTCAGGGACTTGCTTCAAAGCAGGAAAAACGGCATGTTTTGCATCTGGTGTGCAAT
 GTGAAGAGTCTTCAAAAATGCCAGAAATCAACGCCAAGGTGGCTGAATCCACAGAGGAGCCTGCTGGTT
 10 CTAATCGGGGACAGTATCCTGAGGATTCCTCAAGTGATGGTTAAGGCAAAGGGAAGTTCTTCGGAACCT
 TTCTTCCCTGGATGGGAAAACATCTCAAGGCCTGAAGCTGCCAGCAGGCATTCCAAGGCCTGGGTCT
 GGTCTTCTCCGGTTACACACCTATGGGTGGCTTCAGCTTTCCTGGTTCAGCAGATATATGCACGACAGT
 ACTACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTTTGTTCACCAACCAAGTGCAAGAGAT
 ACCTGTGGTCTCTGCACCTGCTCCAGCCCTATTCAACAACAGTTTCCAGCTGAAAACAGCCTGCCAAT
 15 CAGAATGCTGCTCCTCAAGTGGTGTGTAATCCTGGAGCCAATCAAAATTTGCGGATGAATGCACAGGTG
 GCCCTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTGGATGGACCTATTACAGCAGCTACATT
 TTCTGTTTTCTCAGTATCCTCTACTTCTACTCCTCCCTGAGCAGATTCTCATGGTTCATGGGGGCCACC
 GTTGTATGTACCTGCATCACCTTGGGTGGTTTCCATTTAGACCGAGGCCGGTTCCAGAACTTCCAAATG
 ATGGTCTCTCTGACGTTGTAATCAGGACCCCAACAATACTTACAGGAAGGCACTGATCTGAAAC
 TGAAGACCCCAACCACTCCCTCCAGACAGGGATGTACTAGATGGCGAGCAGACCAGCCCTCTTTATG
 20 AGCAGAGCATGGCTTGTCTTCAAGACTTCTTTGCCCTCTCTTCTCCAGAAGGCCCCAGCCATCGCAA
 ACTGATGGTGTGTTGTGCTGTAGCTGTTGGAGGCTTTGACAGGAATGGACTGGATCACCTGATCCAGCTA
 GATTGCCTCTCCTGACATGGCAATGATGAGTTTAAAAAAACAGTGTGGATGATGATGCTTTGTGA
 GCAAGCAAAAGCAGAAACGTGAAGCCGTGATACAAATTGGTGAACAAAAATGCCAAGGCTTCTCATGT
 CTTTATTCTGAAGAGCTTTAATATATACCTCTATGTAGTTTAAATAAGCACTGTACGTAGAAGCCCTAGGT
 25 GTTGCATGTCTATGCTTGAGGAACCTTTCCAAATGTGTGTCTGCATGTGTGTTGTACATAGAAGTCA
 TAGATGCAGAAGTGGTTCTGCTGGTACGATTGATTCCTGTTGGAATGTTTAAATTACACTAAGTGTACT
 ACCTTATATAATCAATGAAATGCTAGACATGTTTAGCAGGACTTTCTAGGAAAGACTTATGTATAAT
 TGCTTTTTAAATGCAGTGCCTTACTTTAACTAAGGGGAACCTTTGCGGAGGTGAAAACCTTTGCTGGGT
 30 TTTCTGTTCAATAAAGTTTTACTATGAAAAAATAAAAAAAAAA

Human HERPUD1 mRNA sequence - var6 (public gi: 12652674)

GAAGTGTCTTGCAGAGATTGCGGGCGGCTGAGACGCGCCCTGCCTGGCACCTAGGAGCGCAGCGGAGCC
 CCGACACCGCCGCGCCGCGCCATGGAGTCCGAGACCGAACCCGAGCCCGTCACGCTCCTGGTGAAGAGCCC
 35 CAACAGCGCCACCGCGACTTGGAGCTGAGTGGCGGACCGCGGCTGGAGTGTGGGCCACCTCAAGGCCAC
 CTGAGCCGCGTCTACCCGAGCGTCCGCGTCCAGAGGACCAGAGGTTAATTTATCTGGGAAGCTGTTGT
 TGGATCACCAATGTCTCAGGGACTTGCTTCAAAGCAGGAAAAACCGCATGTTTGCATCTGGTGTGCAA
 TGTGAAGAGTCTTCAAAAATGCCAGAAATCAACGCCAAGGTGGCTGAATCCACAGAGGAGCCTGCTGGT
 TCTAATCGGGGACAGTATCCTGAGGATTCCTCAAGTGATGGTTAAGGCAAAGGGAAGTTCTTCGGAACC
 40 TTCTTCCCTGGATGGGAAAAACATCTCAAGGCCTGAAGCTGCCAGCAGGCATTCCAAGCCTGGGTCC
 TGGTTTCTCCGGTTACACACCTATGGGTGGCTTCAGCTTTCCTGGTTCAGCAGATATATGCACGACAG
 TACTACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTTTGTTCACCAACCAAGTGCAAGAGA
 TACTGTGGTCTCTGCACCTGCTCCAGCCCTATTCAACAACAGTTTCCAGCTGAAAACAGCCTGCCAA
 TCAGAATGCTGCTCCTCAAGTGGTTGTTAATCCTGGAGCCAATCAAAATTTGCGGATGAATGCACAAGGT
 45 GGCCCTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTTGGATTGGACCTATTACAGCAGTACAT
 TTCTGTTTTTCTCAGTATCCTCTACTTCTACTCCTCCCTGAGCAGATTCTCATGGTTCATGGGGGCCAC
 CGTTGTTATGTACCTGCATCACGTTGGGTGGTTTCCATTTAGACCGAGGCCGGTTCCAGAACTTCCAAAT
 GATGGTCTCTCTGACGTTGTAATCAGGACCCCAACAATACTTACAGGAAGGCACTGATCTGAAA
 CTGAAGACCCCAACCACTCCCTCCAGACAGGGATGTACTAGATGGCGAGCAGACCAGCCCTCTTTAT
 50 GAGCACAGCATGGCTTGTCTTCAAGACTTCTTTCCTCTCTTCTTCCAGAAGGCCCCCAGCCATCGCA
 AACTGATGGTGTGTTGTGCTGTAGCTGTTGGAGGCTTTGACAGGAATGGACTGGATCACCTGATCCAGCT
 AGATTGCCCTCTCCTGACATGGCAATGATGAGTTTAAAAAAACAGTGTGGATGATGATATGCTTTGTG
 AGCAAGCAAAAGCAGAAACGTGAAGCCGTGATACAAATTGGTGAACAAAAATGCCAAGGCTTCTCATG
 TCTTTATTCTGAAGAGCTTTAATATATACCTCTATGTAGTTAATAAGCACTGTACGTAGAAGCCCTAGG
 55 TGTGTCATGTCTATGCTTGAGGAACCTTTCCAAATGTGTGTCTGCATGTGTGTTGTACATAGAAGTC
 ATAGATGCAGAAGTGGTTCTGCTGGTACGATTGATTCTGTTGGAATGTTTAAATTACACTAAGTGTAC
 TACTTTATATAATCAATGAAATGCTAGACATGTTTAGCAGGACTTTCTAGGAAAGACTTATGTATAAT
 TTGCTTTTTAAATGCAGTGCCTTACTTTAACTAAGGGGAACCTTTGCGGAGGTGAAAACCTTTGCTGGG
 TTTCTGTTCAATAAAGTTTTACTATGAATGAAAAAATAAAAAAAAAA

60 Human HERPUD1 mRNA sequence - var7 (public gi: 9711684)

AGAGACGTAAGTGTGCTTGCAGAGATTGCGGGCGGCTGAGACGCGCCCTGCCTGGCACCTAGGAGCGCA
 GCGGAGCCCCGACACCGCCGCGCCGCGCCATGGAGTCCGAGACCGAACCCGAGCCCGTCACGCTCCTGGTG
 AAGAGCCCCAACCGCGCCACCGCGACTTGGAGCTGAGTGGCGACCGCGGCTGGAGTGTGGGCCACCTCA
 AGGCCACCTGAGCCGCGTCTACCCGAGCGTCCGCGTCCAGAGGACCAGAGGTTAATTTATCTGGGAA

5 GCTGTTGTTGGATCACCATGTCTCAGGGACTTGCTTCCAAAGCAGGAAAAACGGCATGTTTTGCATCTG
 GTGTGCAATGTGAAGAGTCCTTCAAAAAATGCCAGAAATCAACGCCAAGGTGGCTGAATCCACAGAGGAGC
 CTGCTGGTTCTAATCGGGGACAGTATCCTGAGGATTCCTCAAGTGATGGTTTAAGGCAAAGGGAAGTTCT
 TCGGAACCTTCTTCCCTGGATGGGAAAAATCTCAAGGCCTGAAGCTGCCAGCAGGCATTCCAAGGC
 10 CTGGGTCTGGTTTCTCCGGTTACACACCTATGGGTGGCTTCAGCTTCTCTGGTTCCAGCAGATATATG
 CACGACAGTACTACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTTTGTTCACCAACCAAGTGC
 ACAAGAGATACCTGTGGTCTCTGCACCTGCTCCAGCCCCTATTACAAACCAGTTTCCAGCTGAAAAACAG
 CCTGCCAATCAGAATGCTGCTCTCAAGTGGTTGTTAATCCTGGAGCCAATCAAAATTTGCGGATGAATG
 CACAAGGTGGCCCTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTTGGATTGGACCTATTTCAGC
 15 AGCTACATTTTCTGTTTTTCTCAGTATCCTCTACTTCTACTCCTCCCTGAGCAGATTCTCATGGTCATG
 GGGGCCACCGTTGTTATGTACCTGCATCAGTTGGGTGGTTTCCATTAGACCCAGGCCGGTTTCAGAACT
 TCCCAAATGATGGTCTCTCTGACGTTGTAATCAGGACCCCAACAATACTTACAGGAAGGCCTGTA
 TCCTGAAACTGAAGACCCCAACACCTCCCTCCAGACAGGGATGTACTAGATGGCGAGCAGACCGCCCT
 TCCTTTATGAGCACAGCATGGCTTGTCTTCAAGACTTTCTTTGCCCTCTCTTCTTCCAGAAGGCCCCAG
 20 CCTCGCAAACTGATGGTGGTTTGTGCTGTAGCTGTTGGAGGCTTTGACAGGAATGGACTGCATCACTGA
 CTCACCTAGATTGCCCTTCTCGGACATGGCAATGATGAGTTTAAAAAAGTGTGGATGATGATATG
 CTTTGTGAGCAAGCAAAAGCAGAAACGTGAAGCCGTGATACAAATTTGGTGAACAAAAATGCCCAAGGC
 TTCTCATGTCTTTATTCTGAAGAGCTTTAATATATACTCTATGTAGTTTAATAAGCACTGTACGTAGAAG
 GCCTTAGGTGTTGCATGCTATGCTTGAGGAACCTTTCCAAATGTGTGTGCTGCATGTGTGTTGTACA
 25 TAGAAGCTATAGATGCAGAAGTGGTTCTGCTGGTACGATTGATTCCCTGTTGGAATGTTTAAATTACACT
 AAGTGTACTACTTTATATAATCAATGAAATTGCTAGACATGTTTTCAGGAGCTTTTCTAGGAAAGACTT
 ATGTATAATTGCTTTTAAAAATGCAGTGCCTTACTTTAAACTAAGGGGAACCTTTGCGGAGGTGAAAACCT
 TTGCTGGGTTTTCTGTTCAATAAAGTTTTACTATGAATGACCCTG

25 Human HERPUD1 mRNA sequence - var8 (public gi: 3005718)

GACGTGAACGGTCGTTGCAGAGATTGCGGGCGGCTGAGACGCCGCTGCTGGCACCTAGGAGCGCAGCG
 GAGCCCCGACACCGCGCGCGCCATGGAGTCCGAGACCGAACCCGAGCCCGTCACGCTCCTGGTGAAG
 AGCCCCAACCGCGCCACCGCGACTTGGAGCTGAGTGGCGACCGCGGCTGGAGTGTGGGCCACCTCAAGG
 30 CCCACCTGAGCCGCGTCTACCCCGAGCGTCCGCGTCCAGAGGACCAGAGGTTAATTTATTCTGGGAAGCT
 GTTGTGGATCACCAATGTCTCAGGGACTTGCTTCCAAAGCAGGAATAACCGGCATGTTTTCATCTGGTG
 TGCAATGTGAAGAGTCCCTTCAAAAATGCCAGAAATCAACGCCAAGGTGGCTGAATCCACAGAGGAGCCTG
 CTCGGTTCTAATCGGGGACAGTATCCTGAGGATTCCCTCAAGTGATGGTTTAAAGGCAAAGGGAAGTTCTTCG
 GAACCTTTCTTCCCTGGATGGGAAAAATCTCAAGGCCTGAAGCTGCCAGCAGGCATTCCAAGGCCTG
 35 GGTCTCGTTTCTCCGGTTACACACCTATGGGTGGCTTCAGCTTCTCTGGTTCCAGCAGATATATGCAC
 GACAGTACTACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTTTGTTCACCAACCAAGTGACACA
 AGAGATACCTGTGGTCTCTGCACCTGCTCCAGCCCCCTATTACAAACCAGTTTCCAGCTGAAAAACAGCCT
 GCCAATCAGAATGCTGCTCTCAAGTGGTTGTTAATCCTGGAGCCAATCAAAATTTGCGGATGAATGCAC
 AAGGTGGCCCTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTTGGATTGGACCTATTTCAGCAGC
 40 TACATTTCTGTTTTTCTCAGTATCCTCTACTTCTCTCTCCCTGAGCAGATTCTCATGGTCATGGGG
 GCCACCGTTGTTATGTACCTGCATCAGTTGGGTGGTTTCCATTAGACCCAGGCCGGTTTCAGAACTTCC
 CAAATGATGGTCTCTCTGACGTTGTAATCAGGACCCCAACAATACTTACAGGAAGGCACTGATCC
 TGAACCTGAAGACCCCAACCACTCCCTCCAGACAGGGATGTACTAGATGGCGAGCAGACAGCCCTCC
 TTTATGAGCACAGCATGGCTTGTCTTCAAGACTTTCTTTGCCCTCTCTTCTTCCAGAAGGCCCCCCAGCCA
 45 TCGCAAACTGATGGTGGTTTGTGCTGTAGCTGTTGGAGGCTTTGACAGGAATGGACTGCATCACTGACTC
 CAGCTAGATTGCCCTCTCTGGACATGGCAATGATGAGTTTTAAAAAACAGTGTGGATGATGATGCTTT
 TTGTGAGCAAGCAAAAGCAGAAACGTGAAGCCGTGATACAAATTTGGTGAACAAAAATGCCCAAGGCTTC
 TCATGTCTTTATTCTGAAGAGCTTTAATATATACTCTATGTAGTTTAATAAGCACTGTACGTAGAAGGCC
 50 TTAGGTGTTGCATGTCTATGCTTGAGGAACCTTTCCAAATGTGTGTGCTGCATGTGTGTTGTACATAG
 AAGTCATAGATGCAGAAGTGGTTCTGCTGGTACGATTGATTCCCTGTTGGAATGTTTAAATTACATAAG
 TGTACTACTTTATATAATCAATGAAATTGCTAGACATGTTTTAGCAGGACTTTTCTAGGAAAGACTTATG
 TATAATTGCTTTTAAAAATGCAGTGCCTTACTTTAAACTAAGGGGAACCTTTGCGGAGGTGAAAACCTTTG
 CTGGGTTTTCTGTTCAATAAAGTTTTACTATGAATGACCCTGAAAAAAAAAAAAAAAAAAAAA

55 Human HERPUD1 mRNA sequence - var9 (public gi: 285960)

CGTGAACGGTCGTTGCAGAGATTGCGGGCGGCTGAGACGCCGCTGCTGGCACCTAGGAGCGCAGCGGA
 GCCCCGACACCGCGCGCGCCATGGAGTCCGAGACCGAACCCGAGCCCGTCACGCTCCTGGTGAAGAG
 60 CCCCACACCGCCACCGCGACTTGGAGCTGAGTGGCGACCGCGGCTGGAGTGTGGGCCACCTCAAGGCC
 CACCTGAGCCGCGTCTACCCGAGCGTCCGCGTCCAGAGGACCAAGAGGTTAATTTATTCTGGGAAGCTGT
 TGTGGATCACCATGTCTCAGGGACTTGCTTCCAAAGCAGGAAAAACCGCATGTTTTCATCTGGTGTG
 CAATGTGAAGAGTCCCTTCAAAAATGCCAGAAATCAACGCCAAGGTGGCTGAATCCACAGAGGAGCCTGCT
 GGTCTAATCGGGGACAGTATCCTGAGGATTCTCAAGTGATGGTTTAAAGGCAAAGGGAAGTTCTTCGGA
 ACCTTTCTTCCCTGGATGGGAAAAATCTCAAGGCCTGAAGCTGCCAGCAGGCATTCCAAGGCCTGGG
 TCCTGGTTTCTCCGGTTACACACCTATGGGTGGCTTCAGCTTCTCTGGTTCCAGCAGATATATGCACGA
 CAGTACTACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTTTGTTCACCAACCAAGTGCACAAG

AGATACCTGTGGTCTCTGCACCTGCTCCAGCCCCTATTACAAACAGTTTCAGCTGAAAACAGCCTGC
 CAATCAGAAATGCTGCTCCTCAAGTGGTTGTTAATCCTGGAGCCAATCAAAATTTGCGGATGAATGCACAA
 GGTGGCCCTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTTGGATTGGACCTATTTCAGCAGCTA
 CATTTCCTGTTTTCTCAGTATCCTCTACTTCTACTCCTCCTGAGCAGATTTCCTCATGGTCATGGGGGC
 5 CACCGTTGTTATGTACCTGCATCACGTTGGGTGGTTCCATTAGACCGAGGCCGGTTCAGAACTTCCCA
 AATGATGGTCTCCTCCTGACGTTGTAAATCAGGACCCCAACAATACTTACAGGAAGGCACGTGATCCTG
 AACTGAAAGACCCCAACCACTCCTCCAGACAGGGATGTACTAGATGGCGAGCAGACCCAGCCCTCCTT
 TATGAGCACAGCATGGCTTGTCTTCAAGACTTTCTTGGCTCTCTTCTTCCAGAAGGCCCCAGCCATC
 10 GCAAACTGATGGTGTGTTGTGCTGTAGCTGTGGAGGCTTTCAGAGGAATGGACTGGATCACTGACTCCA
 GCTAGATTGCCTCTCCTGGACATGGCAATGATGAGTTTTTAAAAAACAGTGTGGATGATGATATGCTTTT
 GTGAGCAAGCAAAAGCAGAAACGTGAAGCCGTGATACAAATGGTGAACAAAAATGCCCAAGGCTTCTC
 ATGTGTTTTATTCTGAAGAGCTTAAATATATACTCTATGTAGTTTAAATAGCACTGTACGTAGAAGGCCTT
 AGGTGTTCATGTCTATGCTTGAGGAACCTTTCCAAATGTGTGTCTGCATGTGTGTTGTACATAGAA
 15 GTCATAGATGCAGAAAGTGGTCTGCTGGTAAGATTGATTCTGTGGAATGTTTAAATTACACTAAGTG
 TACTACTTTATATAATCAATGAAATTGCTAGACATGTTTTAGCAGGACTTTTCTAGGAAAGACTTATGTA
 TAATTGCTTTTAAAAATGCAGTGCTTTACTTTAACTAAGGGGAACCTTTCGCGAGGTGAAAACCTTTGCT
 GGGTTTTCTGTTCAATAAAGTTTTACTATGAATGACCCTG

Human HERPUD1 mRNA sequence - var10 (public gi: 7661869)

20 GACGTGAACGGTCGTGTCAGAGATTGCGGGCGGCTGAGACGCGCCTGCCTGGCACCTAGGAGCGCAGCG
 GAGCCCCGACACCGCCGCGCCGATGGAGTCCGAGACCGAACCCGAGCCGCTCACGCTCCTGGTGAAG
 AGCCCCAACAGCGCCACCGCACTTGGAGCTGAGTGGCGACCGCGGCTGGAGTGTGGGCCACTCAAGG
 CCCACCTGAGCCGCGTCTACCCGAGCGTCCGCGTCCAGAGGACAGAGGTTAATTTATCTCGGAAGCT
 GTTGTGGATCACCAATGTCTCAGGGACTTGCTTCCAAAGCAGGAAAAACGGCATGTTTTCATCTGGTG
 25 TGCAATGTGAAGAGTCTTCAAAAAATGCCAGAAATCAACGCCAAGGTGGCTGAATCCACAGAGGAGCCTG
 CTGGTCTCAATCGGGACAGTATCCTGAGGATTCTCAAGTATGGTTTAAAGGCAAGGGAAGTTCTTCG
 GAACCTTCTTCCCTGGATGGGAAAAACATCTCAAGGCTGAAGCTGCCAGCAGGCATTCCAAGGCCTG
 GGTCTCGTTTCCCGGTTACACACCCCTATGGGTGGCTTCAGCTTCTCTGGTTCCAGCAGATATATGCAC
 30 GACAGTACTACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTTTGTTCACCAACCAAGTGACA
 AGAGATACCTGTGGTCTCTGCACCTGCTCCAGCCCTATTACAAACAGTTTCCAGCTGAAAACAGCCT
 GCCAATCAGAAATGCTGCTCCTCAAGTGGTTGTTAATCCTGGAGCCAATCAAAATTTGCGGATGAATGCAC
 AAGGTGGCCCTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTTGGATTGGACCTATTTCAGCAGC
 TACATTTCTGTTTTCTCAGTATCCTCTACTTCTACTCCTCCTGAGCAGATTCTCATGGTCATGGGG
 35 GCCACCGTGTGTTATGTACCTGCATCAGTTGGGTGGTTCCATTAGACCGAGGCCGGTTTCAAACTTCC
 CAAATGATGGTCTCCTCCTGACGTTGTAAATCAGGACCCCAACAATACTTACAGGAAGGCACTGATCC
 TGAACCTGAAGACCCCAACCACTCCCTCCAGACAGGGATGTACTAGATGGCGAGCAGACAGCCCTCC
 TTTATGAGCACAGCATGGCTTGTCTTCAAGACTTCTTTCCTCTCTTCTTCCAGAAGGCCCCCGAGCCA
 TCGCAAACTGATGGTGTGTTGTGCTGTAGCTGTTGGAGGCTTTGACAGGAATGGACTGGATCACCTGACTC
 40 CAGCTAGATTGCCTCTCCTGGACATGGCAATGATGAGTTTTTAAAAAACAGTGTGGATGATGATATGCTT
 TTGTGAGCAAGCAAAAGCAGAAACGTGAAGCCGTGATACAAATGGTGAACAAAAATGCCCAAGGCTTC
 TCATGTCTTTATTCTGAAGAGCTTAAATATATACTCTATGTAGTTTAAATAGCACTGTACGTAGAAGGCC
 TTAGGTGTTGATGCTATGCTTGAAGAACTTTTCCAAATGTGTGTCTGCATGTGTGTTGTACATAG
 AAGTCATAGATGCAGAAAGTGGTCTGCTGGTACGATTGATTCTGTGGAATGTTTAAATTACACTAAG
 TGTACTACTTTATATAATCAATGAAATTGCTAGACATGTTTTAGCAGGACTTTTCTAGGAAAGACTTATG
 45 TATAATTGCTTTTAAAAATGCAGTGCTTTACTTTAACTAAGGGGAACCTTTCGCGAGGTGAAAACCTTTG
 CTGGGTTTTCTGTTCAATAAAGTTTTACTATGAATGACCCTGAAAAA

Human HERPUD1 Protein sequence - var1 (public gi: 16507802)

50 MESETEPEPVTLVKSPPNRHRLDLSGDRGWSVHLKAHLRVYPERPRPEDQRLIYSGKLLLDHQCLR
 DLLPKEKRVHLVLCNVKSPSKMPEINAKVAESTEEPAGSNRGQYPEDSSSDGLRQREVLRLNSSPGWEN
 ISRHHVGFPPFRPRPVQNFPPNDGPPPDVVNQDPNNLQEGTDPETEDPNHLPDDRDLVDGEQTSFSFMST
 AWLVFKTFFASLLPEGPPAIAN

Human HERPUD1 Protein sequence - var2 (public gi: 10441911)

55 MQYLAATAASGAFVPPPSAQEIPVVSAPAPAPIHNFPAENQPANQNAAPQVVVNPANQNLRMNAQGGP
 IVEEDDEINRDWLDWTYSAAATFSVFLSILYFYSSLSRFLMVMGATVVMYLHHVGVFPFRPRPVQNFNDG
 PPPDVVNQDPNNLQEGTDPETEDPNHLPDDRDLVDGEQTSFSFMSTAWLVFKTFFASLLPEGPPAIAN

Human HERPUD1 Protein sequence - var3 (public gi: 3005723)

60 GHLKAHLRVYPERPRPEDQRLIYSGKLLLDHQCLRDLLPKEKRVHLVLCNVKSPSKMPEINAKVAEST
 EEPAGSNRGQYPEDSSSDGLRQREVLRLNSSPGWENISRPEAAQAFQGLGPGFSGYTPYGLQLSWFQQ
 IYARQYVMQYLAATAASGAFVPPPSAQEIPVVSAPAPAPIHNFPAENQPANQNAAPQVVVNPANQNLR
 MNAQGGPIVEEDDEINRDWLDWTYSAAATFSVFLSILYFYSSLSRFLMVMGATVVMYLHHVGVFPFRPRPV
 QNFNDGPPPDVVNQDPNNLQEGTDPETEDPNHLPDDRDLVDGEQTSFSFMSTAWLVFKTFFASLLPEG

PPAIAN

Human HERPUD1 Protein sequence - var4 (public gi: 7661870)

5 MESETEPEFVTLVVKSPNQHRDLELSGDRGWSVGHKLAHLSRVYPERPRPEDQRLIYSGKLLLDHQCLR
 DLLPKQEKRHVLHLVLCNVKSPSKMPEINAKVAESTEEPAGSNRQYYPEDSSSDGLRQREVLRLNLSPPGWE
 NISRPEAAQAFQGLGPGFSGYTPYQWLQLSWFQQIYARQYMQYLAATAASGAFFVPPPSAQEIIPVVSAP
 APAPIHNPFAENQPANQNAAPQVVNPGANQNLRMNAQGGPIVEEDDEINRDWLDWTYSAATFSVFLSI
 10 LYFYSSLRFLMVMGATVVMYLHHVGFPPFRPRPVQNFNDGPPEDVVDNPDNNLQEGTDPETEDPNHL
 PDRDVLDEGTSPSFMSTAWLVFKTFFASLLPEGPPAIAN

Rat HERPUD1 mRNA sequence (public gi: 16758961)

15 AAGACACCAAGTGTGTTGTGGGGTCGCAGACGGCTGCGTCGCCGCCCGTTCGGGCATCCCTGAGCGCAGT
 CGAGCCTCCAGCGCCGAGACATGAGGACCCGAGCCACAGCCCGAGCCGGTCAAGCTGCTGGTGAAGAGCC
 CCAATCAGCGCCACCGGACTTGGAGCTGAGTGGCGACCGCGGTGGAGTGTGAGTCGCCTCAAGGCCCA
 CCTGAGCCGAGTCTACCCCGAACCGCCGCGCCAGAGGACAGAGGTTAATTTATTCTGGGAAGCTGCTG
 TTGGATCACCAATGTCTCCAAGACTTGCTTCCAAGCAGGAAAGCGACATGTTTTCACCTCGTGTGCA
 ATGTGAGGAGTCCCTCAAAAAGCCAGAAGCCAGCACAAGGGTGTGAGTCCACAGAGCAGCCGGACAA
 20 CACTAGTCAGGCACAGTATCTCTGGGATTCCTCAAGCGATGGCTTACGGGAAAGGGAAGTCCCTTCGGAAC
 CTTCTCCTCTGGATGGGAGAACGTCTCTAGGCCTGAAGCCGTCCAGCAGACTTTCCAAGCCCTCGGGC
 CCGGCTTCTCTGGCTACACACCTACGGGTGGCTGCAGCTCTCTGGTTCAGCAGATCTATGCAAGACA
 GTACTACATGCAATACTTGGCTGCCACTGCTCTCAGGAGCTTTGGCCCTACACCAAGTGCAACAAGAA
 ATACCTGTGGTCTCTACACCGGCTCCCGCCCTATACACAACCAAGTTTCCGGCAGAAAAACAGCCGGCCA
 25 ATCAGAAAGCAGCCGCTCAAGCGGTGTGTAATCCCGGAGCCAAATCAGAACTTGGCGATGAATGCACAAGG
 CGGCCCTCTGGTGAAGAAGATGATGAGATAAACCGAGACTGGTGGATTGGACCTACTCAGCAGCGACA
 TTTTCGGTTTCTCAGCATTCTTACTTCTACTCTCTCCTCAGCAGATTCTCATGGTCTATGGGCGCCA
 CCGTAGTCATGTACCTGCACCACGTGGGTGGTTTCCAATCAGACAGAGGCCAGTTCAGAACTTCCAGA
 TGACGGTCCCCCTCAGGAAGCTGCCAACCAGGACCCCAACAATAACCTCCAGGGAGGTTTGGACCTTGAA
 30 ATGGAAGACCCCAACCGCTCCCGTAGGCCGTGAAGTGTGGACCTGAGCATACCAAGCCCTCGTTCA
 TGAGCACAGCATGGCTAGTCTTCAAGACTTTCTTGGCTCTCTTCTTCCGGAAGGCCCAACAGCCCTAGC
 AAAGTATGGCCCTGTGCTCTGTGCTGGAGGCTTTACAGCTTGGACTGGATCGTCCCTGGCGTGGGA
 CTCGAGAGAGTCATTGAAAACCACAGGATGACGATGTGCTTCTGTGCCAAGCAAAAGCACAACTAAGA
 CATGAAGACCGGTACAAACTGAACAGGGCCCTCATGTGCTTATTCTGAAGAGCTTAAATGTATACTGT
 35 ATGTAGTCTCATAGGCACTGTAAACAGAAGGCCAGGGTCGATGTTCTGCTGAGCACCTCCCGAGAG
 TGTGTGCATGTGTGCGGTACATGGAAGTCATAGACGTGTGTGATGTGTGCTCTACATGGAAGTCATAGA
 TGCAGAAACGGTTCTGCTGGTTGATTTGATTCTGTTGGAATGTTGCAATTACCTAAGTGTACTACTT
 TATATAATCAGTGACTTGTAGACATGTTAGCAGGACTTTCTAGGAGAGACTTATGTATCATTTGCTTT
 TTAAGACGAGTGTCTTACTTACTGAGGGCGGCGACTTGGCACAGGTAAAGCCTTTGCGGGGTTTCTGTT
 40 CAATAAAGTTTTGCTATGAACGACAAAAA

Rat HERPUD1 Protein sequence (public gi: 16758962)

45 MEPEPQPEFVTLVVKSPNQHRDLELSGDRGWSVSRKLAHLSRVYPERPRPEDQRLIYSGKLLLDHQCLQ
 DLLPKQEKRHVLHLVLCNVRSPSKKPEASTKGAESTEQPDNTSQAQYPGDSDDGLREREVLRLNLPSPGWE
 NVSRPEAVQQTQGLGPGFSGYTTYQWLQLSWFQQIYARQYMQYLAATAASGAFFPTPSAQEIIPVVSAP
 APAPIHNPFAENQPANQNAQAQAVVNPANQNLRMNAQGGPLVEEDDEINRDWLDWTYSAATFSVFLSI
 50 LYFYSSLRFLMVMGATVVMYLHHVGFPPFRQRPVQNFDDGPPQEAANQDPNNLQGGLDPEMEDPNRL
 PVGREVLDPHTSPSFMSTAWLVFKTFFASLLPEGPPAIAN

Mouse HERPUD1 mRNA sequence (public gi: 11612514)

55 AAAGACGCCAAGTGTGTTGTGTGGTCTCAGACGGCTGCGTCGCCGCCCGTTCGGGCATCCCTGAGCGCAG
 TCGAGCCGCCAGCGACGACATGGAGCCCGAGCCACAGCCCGAGCCGGTCAAGCTGCTGGTGAAGAGT
 CCAATCAGCGCCACCGGACTTGGAGCTGAGTGGCGACCGCAGTGGAGTGTGAGTCGCCTCAAGGCC
 60 ACCTGAGCCGAGTCTACCCCGAGCGCCCGCTCCAGAGGACAGAGGTTAATTTATTCTGGGAAGCTGCT
 GTTGGATCACCAAGTGTCTCCAAGATTGCTTCCAAGCAGGAAAGCGACATGTTTTCACCTTGTGTGC
 AATGTGAAGAATCCCTCCAAATGCCAGAAACGACCAAAAGGGTGTGTAATCCACAGAGCAGCCGGACA
 ACTCTAATCAGACACAGCATCCTGGGACTCCTCAAGTGATGGTTTACGGCAAAGAGAAGTTCTTCGGAA
 CCTTTCTCCCTCCGGATGGGAGAACATCTCTAGGCCTGAGGCTGTCCAGCAGACTTCCAAGGCTGGGG
 65 CCTGGCTTCTCTGGCTACACAACGTATGGGTGGCTGCAGCTCTCTGGTTCCAGCAGATCTATGCAAGGC
 AGTACTACATGCAATACTTAGCTGCCACTGTGCTACAGGAACCTTTGTCCCGACCAAGTGCAACAAG
 GATACCTGTGTGCTCTACACCTGCTCCGGCTCTATACACAACCAAGTTTCCGGCAGAAAAACAGCCGGCC
 AATCAGAATGCAGCTGCTCAAGCGGTGTCAATCCCGAGCCAAATCAGAACTTGGCGATGAATGCACAAG
 GTGGCCCCCTGTGGAGGAAGATGATGAGATAAACCGAGACTGTTTGGATTGGACCTATTCCGAGCGAC

5 GTTTTCTGTTTTCTCCTCAGCATCCTTTACTTCTACTCCTCGCTGAGCAGATTTCTCATGGTCATGGGTGCC
 ACTGTAGTCATGTACCTGCACCACGTGCGGTGGTTCCGTTCCAGACAGAGGCCAGTTCAGAACTTCCCGG
 ATGATGGTGGTCTCGAGATGCTGCCAACCAGGACCCCAACAATAACCTCCAGGGAGGTATGGACCCAGA
 AATGGAAGACCCCAACCGCTCCCCCAGACCGCGAAGTGCTGGACCTGAGCACACCAGCCCTCGTTT
 10 ATGAGCACAGCATGGCTAGTCTTCAAGACTTCTTTGCCTCTCTTCTCCAGAAGGCCACCAGCCCTAG
 CCAACTGATGGCCCTTGTGCTCTGTGCTGGTGGCTTTGACAGCTCGGACTGGATCGTCTGGCTCCGGCT
 CCTTTTCTCCTCCCTGGCGTGGACTCGACAGAGTCATGAAAACCCACAGGATGACATGTGCTTCTGTGCC
 AAGCAAAAGCACAACTAAGACATGAAGCCGTGGTACAACTGAACAGGGCCCTCATGTCGTTATTCTG
 AAGAGCTTTAATGTATACTGTATGTAGTTTCATAGGCACTGTAAGCAGAAGGCCCGGGTCCGATGTTCT
 15 GCCTGAGCACCTCCCCAGATGTGTGTGCATGTGTGCTGTACATGGAAGTCATAGACGTGTGTGCATGTGT
 GCTCTACATGGAAGTCATAGATGCAGAAACGGTTCTGCTGGTTGATTTGATTCTGTTGGAATGTTCAA
 ATTACACTAAGTGTACTACTTTATATAATCAGTGAATTGCTAGACATGTTAGCAGGACTTTTCTAGGAGA
 GACTTATGTAATGCTTTTAAAATGCAGTGTCTTTCTTTAAACCGAGGGTGGCGACTTGGCAGAGGT
 AAAACCTTTGCCGAGTTTCTGTTCATAAAGTTTGTCTATGAATGACTGT

Mouse HERPUD1 Protein sequence (public gi: 11612515)

20 MEPEPQPEPVTLLVKSPNQRRDLELSGDRSWSVSRLKAHLSRVYPERPRPEDQRLIYSGKLLLDHQCLQ
 DLLPKQEKRHVHLVLCNVKNPSKMPETSTKGAESTE QPDNSNQTHPGDSSSDGLRQREVLRLNLSPSGWE
 NISRPEAVQQTFFQGLGPGFSGYTTYGWLQLSWFQOIYARQYMQYLAATAASGTFVPTPSAQEIPVVSTP
 APAPIHNQFPAENQPANQNAQAQAVVNPGANQNLRMNAQGGPLVEEDDEINRDWLDWTYSAATFSVFLSI
 LYFYSSLSRFLVMGATVVMYLHHVGFPPFRQRPVQNFDDGGPRDAANQDPNNNLQGGMDPEMEDPNRL
 PPDREVLDPEHTSPSFMSTAWLVFKTFFASLLPEGPPALAN

25 INCORPORATION BY REFERENCE

All publications and patents mentioned herein are hereby incorporated by
 reference in their entirety as if each individual publication or patent was specifically
 and individually indicated to be incorporated by reference. In case of conflict, the
 present application, including any definitions herein, will control.

30

EQUIVALENTS

35 While specific embodiments of the subject applications have been discussed,
 the above specification is illustrative and not restrictive. Many variations of the
 applications will become apparent to those skilled in the art upon review of this
 specification and the claims below. The full scope of the applications should be
 determined by reference to the claims, along with their full scope of equivalents, and
 the specification, along with such variations.

What Is Claimed:

1. An isolated, purified or recombinant complex comprising a POSH polypeptide and a POSH-AP.
- 5 2. The complex of claim 1, wherein the POSH-AP comprises a polypeptide selected from the group consisting of: selected from the group consisting of: an UNC84, an MSTP28 and an HERPUD1.
3. An isolated, purified or recombinant complex comprising
 - 10 a) a polypeptide comprising a domain that is at least 90% identical to a POSH SH3 domain; and
 - b) a POSH-AP comprising a polypeptide selected from the group consisting of: an UNC84, an MSTP28 and an HERPUD1.
- 15 4. A method of identifying an antiviral or anti-apoptotic agent, the method comprising identifying a test agent that disrupts a complex of any of claims 1-3.
- 4A. A method of identifying an agent to treat a neurological disorder, the method comprising identifying a test agent that disrupts a complex of any of claims 1-3.
- 20 4B. The method of claim 4A, wherein the neurological disorder is a CNS disorder.
- 4C. The method of claim 4B, wherein the CNS disorder is selected from the group consisting of: Alzheimer's disease, cerebral vascular disease and schizophrenia.
- 25 5. A method for identifying an antiviral or antiapoptotic agent comprising:
 - a) providing a POSH-AP polypeptide and a test agent; and
 - b) identifying a test agent that binds to the POSH-AP polypeptide.
- 30 5A. A method for identifying an agent to treat a neurological disorder, comprising:
 - a) providing a POSH-AP polypeptide and a test agent; and
 - b) identifying a test agent that binds to the POSH-AP polypeptide.

- 5B. The method of claim 5A, wherein the neurological disorder is a CNS disorder.
- 5C. The method of claim 5B, wherein the CNS disorder is selected from the group
5 consisting of: Alzheimer's disease, cerebral vascular disease and schizophrenia.
6. The method of claim 5 or 5A, wherein the POSH-AP is selected from the group
consisting of: an UNC84, an MSTP28 and an HERPUD1.
- 10 7. A method of identifying an agent with anti-apoptotic or anti-viral activity, the
method comprising:
a) contacting a POSH-AP polypeptide with a test agent, and
b) identifying a test agent that modulates a POSH-AP activity.
- 15 7A. A method of identifying an agent with activity against the progression of a
neurological disorder, the method comprising:
a) contacting a POSH-AP polypeptide with a test agent, and
b) identifying a test agent that modulates a POSH-AP activity.
- 20 7B. The method of claim 7A, wherein the neurological disorder is a CNS disorder.
- 7C. The method of claim 7B, wherein the CNS disorder is selected from the group
consisting of: Alzheimer's disease, cerebral vascular disease and schizophrenia.
- 25 8. The method of claim 7 or 7A, wherein the POSH-AP is selected from the group
consisting of: an UNC84, an MSTP28 and an HERPUD1.
9. A method for identifying an antiviral or anti-apoptotic agent comprising:
providing a POSH-AP and a test agent; and
30 identifying a test agent that interacts with the POSH-AP.
- 9A. A method for identifying an agent to treat a neurological disorder, comprising:

providing a POSH-AP and a test agent; and
identifying a test agent that interacts with the POSH-AP.

- 9B. The method of claim 9A, wherein the neurological disorder is a CNS disorder.
- 5 9C. The method of claim 9B, wherein the CNS disorder is selected from the group consisting of: Alzheimer's disease, cerebral vascular disease and schizophrenia.
- 10 10. The method of claim 9 or 9A, wherein the POSH-AP is selected from the group consisting of: an UNC84, an MSTP28 and an HERPUD1.
- 15 11. An isolated antibody, or fragment thereof, specifically immunoreactive with an epitope of a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID No: 26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, and SEQ ID NO:30, which antibody disrupts the interaction between a polypeptide of SEQ ID NO: 2 and a POSH-AP.
- 20 12. A method of inhibiting viral infection comprising administering an agent to a subject in need thereof wherein said agent inhibits the interaction between a POSH polypeptide and a POSH-AP.
- 25 12A. A method of inhibiting the progression of a neurological disorder, comprising administering an agent to a subject in need thereof wherein said agent inhibits the interaction between a POSH polypeptide and a POSH-AP.
- 30 12B. The method of claim 12A, wherein the neurological disorder is a CNS disorder.
- 12C. The method of claim 12B, wherein the CNS disorder is selected from the group consisting of: Alzheimer's disease, cerebral vascular disease and schizophrenia.

13. The method of claim 12 or 12A, wherein the POSH-AP is selected from the group consisting of: an UNC84, an MSTP28 and an HERPUD1.

ABSTRACT

The application discloses novel polypeptides and nucleic acids involved in a
5 variety of biological processes, including viral reproduction. Related methods and
compositions are also described.

Figure 1: Human POSH Coding Sequence (SEQ ID NO:1)

ATGGATGAATCAGCCTTGTTGGATCTTTTGGAGTGTCCGGTGTGTCTAGAGCGCCTTGATGCTTCTGCGA
 AGGTCTTGCCCTTGCCAGCATACGTTTGGCAAGCGATGTTTGGCTGGGGATCGTAGGTTCTCGAAATGAACT
 CAGATGTCCCGAGTGCAGGACTCTTGTTGGCTCGGGTGTGAGGAGCTTCCAGTAACATCTTGCTGGTC
 5 AGACTTCTGGATGGCATCAAACAGAGGCCCTGGAAACCTGGTCTGGTGGGGGAAGTGGGACCAACTGCA
 CAAATGCATTAAAGGTCTCAGAGCAGCACTGTGGCTAATTGTAGCTCAAAGATCTGCAGAGCTCCAGGG
 CGGACAGCAGCCTCGGGTGCAATCCTGGAGCCCCCAGTGAGGGGTATACCTCAGTTACCATGTGCCAAA
 GCGTTATACAACATGAAGGAAAAGAGCCTGGAGACCTTAAATTGAGCAAAGGCGACATCATCATTTTGC
 10 GAAGACAAGTGGATGAAAATTGGTACCATGGGGAAGTCAATGGAATCCATGGCTTTTCCCCACCAACTT
 TGTGCAGATTATTAACCGTTACCTCAGCCCCACCTCAGTGCAAAGCACTTTATGACTTTGAAGTGAAA
 GACAAGGAAGCAGACAAAGATTGCCTTCCATTTGCAAAGGATGATGTTCTGACTGTGATCCGAAGATGG
 ATGAAAACTGGGCTGAAGGAATGCTGGCAGACAAAATAGGAATATTTCCAATTTTATATGTTGAGTTTAA
 CTCGGCTGCTAAGCAGCTGATAGAAATGGGATAAGCTCCTGTGCCAGGAGTTGATGCTGGAGAATGTTCC
 15 TCGGCAGCAGCCAGAGCAGCACTGCCCCAAAGCACTCCGACACCAAGAAGAACACCAAAAGCGGCACT
 CCTTCACTTCCCTCACTATGGCCAAAGTCTCCAGGCATCCAGAACCGCCACTCCATGGAGATCAG
 CCCCCCTGTCTCATCAGCTCCAGCAACCCCACTGCTGCTGCACGGATCAGCGAGCTGTCTGGGCTCTCC
 TGCAGTGCCCCCTTCTCAGGTTTATATAAGTACCACCGGTTAATTGTGACCCCGCCCCAACGAGCCAG
 TGACAACTGGCCCCCTCGTTTACTTTCCCATCAGATGTTCCCTACCAAGCTGCCCTTGGAACTTTGAATCC
 20 TCCTCTTCCACCAACCCCTCTCCTGGCTGCCACTGTCTTGCCTCCACACCACAGGCGCCACCGCCGCC
 GCTGTGCTGTGGAATGGGACCGAGGCCATGGCAGGATCCACTGACCAGATTGCACATTTACGGCCGC
 AGACTCGCCCCAGTGTGTATGTTGCTATATCCATACACTCCTCGGAAAGAGGATGAACAGAGCTGAG
 AAAAGGGGAGATGTTTTAGTGTTTGGAGCGCTGCCAGGATGGCTGGTTCAAAGGGACATCCATGCATACC
 AGCAAGATAGGGGTTTTCCCTGGCAATTATGTGGCACCAGTCAAGGGCGGTGACAAATGCTTCCCAAG
 25 CTAAAGTCCCTATGTCTACAGCTGGCCAGACAAAGTCCGGGAGTGACCATGGTCAGTCTTCCACGGCAGG
 AGGGCCTGCCCAGAAGCTCCAGGGAATGGCGTGGCTGGGAGTCCAGTGTGTCCCCGAGCTGTGGTA
 TCAGCAGCTCACATCCAGACAAGTCTCAGGCTAAGGTCTTGTGACATGACGGGGCAATGACAGTCA
 ACCAGGCCCGCAATGCTGTGAGGACAGTTGCAGCGCACAAACAGGAACGCCCCACGGCAGCAGTGACACC
 CATCCAGGTACAGAATGCCGCCGCCCTCAGCCCTGCATCTGTGGGCTGTCCCATCACTCGCTGGCCTCC
 30 CCACAACCTGCGCCTCTGATGCCAGGCTCAGCCACGCACACTGTGCCATCAGTATCAGTCGAGCCAGTG
 CCCCTCTGGCCTGTGCAGCAGCTGCTCCACTGACTTCCCCAAGCATCACCAGTGTCTCTGGAGGCTGA
 GCCCAGTGGCCGGATAGTGACCGTTCTCCCTGGACTCCCCACATCTCCTGACAGTGTTCATCAGCTTGT
 GGGACAGTTTACGAACCAAACAGACAAGGATAGCAAAAAAGAAAAAGGGTTTTGTTGAAGTTGCTTT
 CTGGCGCCTCCACTAAACGGAAGCCCGCGTGTCTCCTCCAGCATCGCCCAACCTAGAAAGTGGAGCTGGG
 35 CAGTGACAGACTTCTCTCCAGGGAGCGGTGGGGCCCCAACTGCCACCAGGAGGTGGCCATGGCAGGGCA
 GGCTCCTGCCCCGTGGACGGGACGGACCGGTACAGACTGCAGTGGCAGGAGCAGCCCTGGCCACGGATG
 CTTTTCATAGGAAGGCAAGTTCCCTGGACTCCGCAAGTCCCATCGCTCCACCTCCTCGCCAGGCCTGTTT
 CTCCTTGGGTCCTGTCTTGAATGAGTCTAGACCTGTCGTTTGTGAAAGGCACAGGGTGGTGGTTTCTTAT
 CCTCCTCAGAGTGAGGCAGAACTTGAACCTAAAGAAGGAGATATTGTGTTTGTTCATAAAAAACGAGAGG
 40 ATGGCTGGTTCAAAGGCACATTACAACGTAATGGGAAAACCTGGCCTTTTCCAGGAAGCTTTGTGAAAA
 CATATGA

Figure 2: Human POSH Amino Acid Sequence (SEQ ID NO:2)

MDESALLDLLECPVCLERLDASAKVLPQCQHTFCKRCLLGIVGSRNELRCPECRTLIVGSGVEELPSNILLV
5 RLLDGIKQRPWKPGPGGGSGTNCNLRSSQSTVANCSSKDLQSSQGGQQPRVQSWSPVVRGIPQLPCAK
ALYNYEGKEPGDLKFSKGDIIILRRQVDENWYHGEVNGIHGFPPTNFVQIIKPLPQPPPPQCKALYDFEVK
DKEADKDCLPFAKDDVLTIVIRRVDENWAEGLADKIGIFPISYVEFNAAKQLIEWDKPPVPGVDAGECS
SAAQSSSTAPKHSDDTKNTKKRHSFTSLTMANKSSQASQNRHSMEISPPVLISSSNPTAAARISELSGLS
CSAPSQVHI STTGLIVTPPPSSPVTGSPSTFPSPDVPYQAALGTLNPPPLPPPLLAATVLA STPPGATAA
10 AAAAGMGPRPMAGSTDQIAHLRPQTRPSVYVAIYPTPRKEDELELRKGEMFLVFERCQDGFKGTSMHT
SKIGVFPNGYVAPVTRAVTNASQAKVPMSTAGQTSRGVTMVSPSTAGGPAQKLQNGVAGSPSVVPAAVV
SAAHIQTSPQAKVLLHMTGQMTVNQARNVRTVAAHNQERPTAAVTPIQVQNAAGLSPASVGLSHSLAS
PQPAPLMPGSATHTAASISRSASPLACAAAAPLTSPSITSASLEAEPGGRIVTVLPGLPTS PDSASSAC
GNSSATKPKDKSKKEKGLLKLKLSGASTKRKPRVSPASPTLEVELGSAELPLQGA VGPELPPGGGHGRA
15 GSCPVDGDGPVTTAVAGAALAQDAFHRKASSLDSAVPIAPPPRQACSSLG PVLNESRPVVCERHRVVVSY
PPQSEAELELKEGDIVFVHKREDGWFKGTIQRNGKTGLFPGSFVENI

Figure 3: Human POSH cDNA Sequence (SEQ ID NO:3)

CTGAGAGACACTGCGAGCGGCGAGCGCGTGGGGCCGCATCTGCATCAGCCGCCGAGCCGCTGCGGGGC
 5 CGCGAACAAAGAGGAGGAGCCGAGGCGCGAGAGCAAAGTCTGAAATGGATGTTACATGAGTCATTTTAAG
 GGATGCACACAACCTATGAACATTTCTGAAGATTTTTTCTCAGTAAAGTAGATAAAGATGGATGAATCAGC
 CTTGTTGGATCTTTTGGAGTGTCCGGTGTGTCTAGAGCGCCTTGATGCTTCTGCGAAGGCTTGCCTTGC
 CAGCATACGTTTGGCAAGCGATGTTTGTGGGGATCGTAGGTTCTCGAAATGAACTCAGATGTCCCGAGT
 GCAGGACTCTTGTGGCTCGGGTGTGAGGAGCTTCCAGTAACATCTTGCTGGTCAGACTTCTGGATGG
 10 CATCAACAGAGGCCCTTGGAAACCTGGTCTGTGGGGGAAGTGGGACCAACTGCACAAATGCATTAAAG
 TCTCAGAGCAGCACTGTGGCTAATTGTAGTCAAAAGATCTGCAGAGCTCCAGGGCGGACAGCAGCCTC
 GGGTGCAATCTGGAGCCCCCAGTGAGGGGTATACCTCAGTTACCATGTGCCAAAGCGTTATACAATA
 TGAAGGAAAAGAGCCTGGAGACCTTAAATTAGCAAAGGCGACATCATCATTTTGCAGAGACAAGTGGAT
 GAAATTGGTACCATGGGAAGTCAATGGAATCCATGGCTTTTCCCACTAATTTGTGCAAGATTATTA
 15 AACCGTTACCTCAGCCCCCAGTGCAGGACACTTTATGACTTTGAAAGTGAAGAACAAGGAAGCAGA
 CAAAGATTGCTTCCATTTGCAAAGGATGATGTTCTGACTGTGATCCGAAGAGTGGATGAAACTGGGCT
 GAAGGAATGCTGGCAGACAAAATAGGAATATTTCCAATTTTATATGTTGAGTTTAACTCGGCTGCTAAGC
 AGCTGATAGAATGGGATAAGCCTCTGTGCCAGGAGTTGATGCTGGAGAATGTTCTCGGACAGCCCA
 GAGCAGCACTGCCCCAAAGCACTCCGACACCAAGAAGAACCAAAAAGCGGCACTCCTTCACTTCCCTC
 20 ACTATGGCCAAAGTCTCCAGGCATCCAGAACCGCCACTCCATGGAGATCAGCCCCCTGTCTCTCA
 TCAGTCCAGCAACCCCACTGCTGCTGCAGGATCAGCGAGCTGTCTGGGCTCTCTGCAAGTCCCTTC
 TCAGGTTTCAATAAGTACCACCGGTTAATTGTGACCCCGCCCCAAGCAGCCAGTGAACCTGGCCCC
 TCGTTTACTTTCCCATCAGATGTTCCCTACCAAGCTGCCCTTGGAACTTTGAATCCTCCTCTCCACCAC
 CCCCCTCTCTGGCTGCCACTGTCTTGCCTCCACACCAGGCGCCACCGCCGCGCTGTCTGTCTGG
 25 AATGGGACGAGGCCCATGGCAGGATCCACTGACACGATTTGACATTTACGGCCGAGACTCGCCCACT
 GTGTATGTTGCTATATATCCATACATCTCTCGGAAAGAGGATGAACTAGAGCTGAGAAAAGGGAGATGT
 TTTTAGTGTTTGGCGCTGCCAGGATGGCTGGTTCAAAGGGACATCCATGCATACCAGCAAGATAGGGGT
 TTTCCCTGGCAATTATGTGGCACCAGTCACAAGGGCGGTGACAAATGCTTCCCAAGCTAAAGTCCCTATG
 30 TCTACAGCTGGCCAGACAAGTCCGGGAGTGACCATGGTCAGTCTTCCACGGCAGGAGGCTGCCAGAG
 AGCTCCAGGGAAATGGCGTGGCTGGGAGTCCAGTGTGTCTCCCGCAGCTGTGGTATCAGCAGCTCACAT
 CCAGACAAGTCTCAGGCTAAGGTCTTGTTCACATGACGGGGCAAATGACAGTCAACCAGGCCCCGCAAT
 GCTGTGAGGACAGTTGCAGCGCACAAACAGGAACGCCACGGCAGCAGTGAACCCATCCAGGTACAGA
 ATGCGCGCGGCTCAGCCCTGCATCTGTGGGCTGTCCCATCACTCGCTGGCCTCCCCACAACCTGCGCC
 35 TCTGATGCCAGGCTCAGCCACGCACACTGCTGCCATCAGTATCAGTCGAGCCAGTGCCTCTGCGCTGT
 GCAGCAGCTGCTCCACTGACTTCCCCAAGCATCACCAGTGCTTCTCTGGAGGCTGAGCCAGTGGCCGGA
 TAGTGACCGTTCTCCCTGGACTCCCCACATCTCTGACAGTGCTTCATCAGCTTGTGGGAACAGTTTCAGC
 AACCAAAACCAGACAAGGATAGCAAAAAAGAAAAAGGGTTTGTGAAGTTGCTTCTGGCGCCTCCACT
 AAACGGAGAGCCCGCGTGTCTCTCCAGCATCGCCACCTAGAGTGGAGCTGGGCAGTGCAGAGCTTC
 40 CTCTCCAGGGAGCGGTGGGGCCGAACTGCCACCAGGAGTGGCCATGGCAGGGCAGGCTTCCCTGCT
 GGACGGGACGGACCGGTACGACTGCAGTGGCAGGAGCAGCCCTGGCCAGGATGCTTTTCATAGGAAG
 GCAAGTTCCTGGACTCCGCAAGTCCCATCGCTCCACCTCTCGCCAGGCTGTTCCTCCCTGGGTCTCTG
 TCTTGAATGAGTCTAGACCTGTGCTTTGTGAAGGACAGGGTGGTGGTTTCTATCCTCTCAGAGTGA
 45 GGCAGAACTTGAACCTAAAGAAGGAGATATTGTGTTTGTTCATAAAAAACGAGAGGATGGCTGGTTCAA
 GGCACATTACAACGTAATGGGAAAACCTGGCCTTTTCCAGGAAGCTTGTGGAAAAATATGAGGAGACT
 GACACTGAAGAAGCTTAAATCACTTACACAACAAAGTAGCACAAAGCAGTTAAACAGAAAGAGCACAT
 TTGTGGACTTCCAGATGGTCAGGAGATGAGCAAGGATTGGTATGTGACTCTGATGCCCCAGCACAGTTA
 50 CCCAGCGAGCAGAGTGAAGAAGATGTTTGTGTGGGTTTGTGTAGTCTGGATTGGATGTATAAGGTGTG
 CCTGTACTGTCTGATTTACTACACAGAGAACTTTTTTTTTTTTTTAAGATATATGACTAAAATGGACA
 ATTGTTTACAAGGCTTAACATAATTTATTTGCTTTTTTAAACTTGAACTTTTCGTATAATAGATACGTTCT
 55 TTGGATTATGATTTTAAGAAATTAATTAATTTATGAAATGATAGGTAAGGAGAAGCTGGATTATCTCTGT
 TGAGAGCAAGAGATTGTTTTGACATAGAGTGAATGCATTTTCCCTCTCTCTCTCTCTCTCTCTCTCTCT
 ATTTTGGGTTATGTTTTGCTTCTTTAAGATAGAAATCCAGTCTCTTAATTTGGTTTTCTTCTTTGGGA
 AACCAACATACAAATGAATCAGTATCAATTAGGGCTCGGGTAGAGAGACAGAACTTGAGAGAAGAGA
 60 AGTTAGTGATTTCCCTCTCTTTCTAGTTTGGTAGGAATCACCTGAAGACCTAGTCTCTCAATTAATTTGTG
 TGGGTTTTTAATTTCTTAGAATGAAGTGAAGTGAACCAATGAGAAAGAAACAGCAACACCTTGAACAA
 AATGTATTTAGAAATATATTTAGTTTTATAGCAGAGCAGCTCAATTTGTTTGGTTGGAAAGTAGGGGAAA
 TTGAAGTTGTAGTCACTGTCTGAGAATGGCTATGAAGCGTCATTTTCAATTTTACCCCACTGACCTGCA
 65 TGCCAGGACACAAGTAAACATTTGTGAGATAGTGGTGGTAAGTGAATGCACTCGTGTAAAGTCAAAGGC
 TATAAGAAACACTGTGAAAGTTTATATTCATCCATTGTGATTTCTTCCCCACGCTTGTGATGTATTACT
 GGATTCCCACAGTAATATAGACTGTGCATGGTGTGTATATTTCATTGCGATTTCCTGTTAAGATGAGTTT
 GTACTCAGAAATTGACCAATTGAGAGGTGTAATAAATGAAACAGTGTCTCTTCTCTACCCCAAGCCACTA
 CTGACCAAGGTCTCTTCACTGCACTCGCTCCCTCTCTGGCTAAGGCATGCATTAGCCACTACACAAGTCA
 TTAGTGAAAGTGGTCTTTATGTCTCTCCAGCAGACAGACATCAAGGATGAGTTAAACAGGAGACTACTC
 CTGTGACTGTGGAGCTCTGGAAGGCTTGGTGGGAGTGAATTTGCCACACCTTACAATTTGTGGCAGGATC
 CAGAAGAGCCTGTCTTTTATATCCATTCTTGTATGTCATTGGCCTCTCCACCGATTTCATTACGGTGC
 CACGCACTCATGGATCTGGGTAGTCCGGAACAAAGAGGGAAGACAGCCTGGTAATGAATAAGATCC

TTACCACAGTTTTCTCATGGGAAATACATAATAAACCCCTTTCATCTTTTTTTTTTCTTTAAGAATTAA
AACTGGGAAATAGAAACATGAACTGAAAAGTCTTGCAATGACAAGAGGTTTCATGGTCTTAAAAAGATAC
TTTATATGGTTGAAGATGAAATCATTCTAAATTAACCTTTTTTTTTAAAAAAAACAATGTATATTATGT
5 TCCTGTGTGTTGAATTTAAAAAAAATACTTACTTGGATATTTCATGTAATATATAAAGGTTTGGTG
AAATGAACCTTAGTTAGGAAAAAGCTGGCATCAGCTTTCATCTGTGTAAGTTGACACCAATGTGTCATAA
TATTCTTTATTTTGGGAAATTAGTGTATTTTATAAAAAATTTAAAAAGAAAAAGACTACTACAGGTTAA
GATAATTTTTTACCCTGTCTTTCTCCATATTTTAAGCTATGTGATTGAAGTACCTCTGTTCATAGTTTC
10 CTGGTATAAAGTTGGTTAAAAATTCATCTGTTAATAGATCATTAGGTAATATAATGTATGGGTTTTCTAT
TGGTTTTTTGCAGACAGTAGAGGGAGATTTTGTAACAAGGGCTTGTTACACAGTGATATGGTAATGATAA
AATTGCAATTTATCACTCCTTTTCATGTTAATAATTTGAGGACTGGATAAAAGGTTTCAAGATTAAATTT
TGATGTTCAAACCTTGT

Figure 4: 5' cDNA fragment of human POSH (public gi:10432611; SEQ ID NO:4)

```

ctgagagacactgcgagcggcgagcgcggtggggccgcacatctgcatcagccgcccagccgctgcggggc
5 cgcgaaacaaagaggaggagccgaggcgcgagagcaaaagtctgaaatggatgttacatgagtcattttaag
gatgcacacaactatgaacatttctgaagatttttctcagtaaaagtagataaagatggatgaatcagcc
ttgttggatcctttggagtggtccggtgtgtctagagcgcccttgatgcttctcgaaaggtcttgccttgc
agcatacgttttgaagcgatgtttgctggggatcgtaggttctcgaaatgaactcagatgtcccagtg
caggactcctgttggctcgggtgtcgaggagcttcccagtaacatcttgctgggtcagacttctggatggc
10 atcaaacagagccttggaaacctggctcctgggtgggggaagtgggaccaactgcacaaatgcattaaagt
ctcagagcagcactgtggctaattgttagctcaaaagatctgcagagctcccaggcggaacagcagcctcg
gggtgcaatcctggagccccccagtgaggggtatacctcagttaccatgtgccaaagcgttatacaactat
gaaggaaaagagcctggagaccttaaattcagcaaaaggcgacatcatctttcggaagacaagtggatg
aaaattgggtaccatggggaagtcaatgggaatccatggcttttccccaccaactttgtgcagattattaa
15 accggtaccctcagccccacacctcagtgcaaacgacctttatgactttgaagtgaagacaagggaagcagac
aaagattgcccttccatttgcaaaaggatgatgttctgactgtgatccgaagagtggatgaaaactgggctg
aaggaaatgctggcagacaaaataggaatatttccaatttcatatgttgagtttaactcgggtgctaagca
gctgatagaatgggataagcctcctgtgcccaggagttgatgctggagaatgttcctcggcagcagcccag
agcagcactgccccaaagcactccgacaccaagaagaacacccaaaagcggcactccttcacttcctca
20 ctatggccaacaagtctcccaggcatcccagaaccgccaactccatggagatcagccccctgtcctcat
cagctccagcaacccactgctgctgcacggatcagcgagctgtctgggctctcctgcagtgccttctc
caggttcatataagtaccacgggttaattgtgaccccgcccccaagcagcccagtgacaactggccctc
cgtttactttcccatcagatgttcctaccaagctgcccttggaaactttgaatcctcctctccaccacc
ccctctcctgggtgccactgtccttgccctccacaccaccaggcgccaccgcccgcgctgctgctgtgga
25 atgggaccgaggcccatggcaggatccactgaccagatgacacatttacggccgcagactcgccccagtg
tgtatgttgctatatatccatacactcctcggaagaggatgaactagagctgagaaaaggggagatgtt
tttagtgtttgagcgctgccaggatggctgggttcaaagggacatccatgcataccagcaagataggggtt
ttccctggcaattatgtggcaccagtcaaaagggcggtgacaaatgcttcccaagctaaagtccctatgt
ctacagctggccagacaagtggggagtgaccatggtcagtcttccacggcaggagggcctgccagaa
30 gctccaggggaaatggcgtggctgggagteccagtggtgtccccgcagctgtgggtatcagcagctcacatc
cagacaagtctcagggttaaggctctgtgtgacatgacggggcaaatgacagtcaaccaggcccgcaatg
ctgtgaggacagttgcagcgacacaaccaggaaacgccccacggcagcagtgacacccatccaggtaacagaa
tgccgcccgcctcagcctgcacatctgtgggcctgtcccatcactcgctggcctccccacaacctggcct
ctgatgccaggctcagccacgcacactgctgccatcagtatcagtcgagccagtgcctcctggcctgtg
35 cagcagctgctccactgacttcccaagcatcaccagtgtctctcggaggctgagcccagtgggccgat
agtgaccgttctcctggactccccacatctcctgacagtgcttcatcagcttgtgggaaagttcagca
accaaaccagacaaggatagc

```

Figure 5: N terminus protein fragment of hPOSH (public gi:10432612; SEQ ID NO:5)

5 MDESALLDLECPVCLERLDASAKVLPCQHTFCKRCLLGIVGSRNELRCPECRTLVGSGVEELPSNILLV
RLLDGIKQRPWKPGPGGGSGTNCNALRSQSSTVANCSSKDLQSSQGGQQPRVQSWSPFVRGIPQLPCAK
ALYNYEGKEPGDLKFSKGDIIILRRQVDENWYHGEVNGIHGFFPTNFVQIIKPLPQPPPQCKALYDFEVK
DKEADKDCLEPAKDDVLTVIRRVDENWAEGLADKIGIFPISYVEFNAAKQLIEWDKPPVPGVDAGECS
SAAAQSSTAPKHSSTKNTKKRHSFTSLTMANKSSQASQNRHSMEISPPVLISSSNPTAAARISELSGLS
10 CSAPSQVHISTTGLIVTPPPSSPVTTGPSFTFPDVPYQAALGTLPPLPPPPPLAATVLAATPPGATAA
AAAAGMGPRPMAGSTDQIAHLRPQTRPSVYVAIYPYTPRKEDELELRKGEMFLVFERCQDGWFKGTSMHT
SKIGVFPNGYVAPVTRAVTNASQAKVPMSTAGQTSRGVTMVSPSTAGGPAQKLQNGVAGSPSVVPAAVV
SAAHIQTSPOAKVLLHMTGQMTVNQARNVARTVAAHNQERPTAAVTPIQVQNAAGLSPASVGLSHHSLAS
PQPAPLMPGSATHTAAISISRASAPLACAAAAPLTPSPITSASLEAEPGRIVTVLPGLTPSPDSASSAC
GNSSATKPKDKS

Figure 6: 3' mRNA fragment of hPOSH (public gi:7959248; SEQ ID NO:6)

5 atttcatatgttgagtttaactcggctgctaagcagctgatagaatgggataagcctcctgtgccaggag
 ttgatgctggagaatgttccctcggcagcagccagagcagcactgccccaaagcactccgacaccaagaa
 10 gaacacaaaaagcggcactccttcacttccctcactatggccaacaagtccctccaggcatccagaac
 cgccactccatggagatcagccccctgtcctcatcagctccagcaacccccactgctgtgcacggatca
 gcgagctgtctgggctcctcctgcagtgccccctctcagggttcataaagtaccacgggttaattgtgac
 cccgcccccaagcagccagtgacaactggccccctcgtttactttcccatcagatgttccctaccaagct
 15 gcccttggaaactttgaatcctcctcttccacacccccctcctcctggctgccactgtccttgcctccacac
 caccaggcgccaccgcgctgctgctgctgctggaatgggacggagcccatggcaggatccactgacca
 gattgacacatttacggcgcgagactcggccccagtgtgtatgttgctatatatccatacactcctcggaaa
 gaggatgaactagagctgagaaaagggagatgtttttagtgtttgagcgtgccaggatggctggttca
 aagggacatccatgcataccagcaagataggggttttccctggcaattatgtggcaccagtacaaagggc
 20 ggtgacaaatgcttcccaagctaaagtccctatgtctacagctggccagacaagtgcgggagtgacacat
 gtcagtctctccacggcaggagggtcctgccagaagctccagggaatggcggtggctgggagctccagtg
 ttgtccccgcagctgtggtatcagcagctcacatccagacaagtccctcaggctaaggtcttgttgacat
 gacggggcaaatgacagtcacacaggcccgcaatgtgtgaggacagttgcagcgcacaaccagggaacgc
 cccacggcagcagtgacacccatccaggtacagaatgccgcggcctcagccctgcactctgtgggctgt
 25 cccatcactcgtggcctccccacaactgcgcctctgatgccaggctcagccacgcacactgtgccat
 cagtatcagtcgagccagtgcctcctgtgctgtgcagcagctgtccactgacttccccaaagcatcacc
 agtgcttctctggaggctgagccccagtggccggatagtgcagcttctcctggactccccacatctcctg
 acagtgttctcagcttgtgggaacagttcagcaaccaaaccagacaaggatagcaaaaaagaaaaaa
 ggggtttgttgaaagtgtcttctggcgccctccactaaacgggaagccccgcgtgtctcctccagcatcgccc
 30 accctagaagtggagctgggcagtgagagcttccctcctcagggaagcgggtggggcccgaaactgccaccag
 gaggtggccatggcagggcaggtcctcctgctgtggacggggacggacgggtcacgactgcagtgccagg
 agcagccctggccaggatgtcttcataggaaggcaagtccctggactccgcagttcccatcgctcca
 cctcctcgccaggcctgttctcctgggtcctgtcttgaatgagcttagacctgtcgtttgtgaaagggc
 acaggggtgggtgttccctatcctcctcagagtgaggcagaacttgaacttaagaaggagatattgtgtt
 35 tgttcataaaaaacgagaggatgggtgttcaaggacattacaacgtaattgggaaaactggcctttc
 ccagggaagctttgtggaaaacatagaggagactgacactgaagaagcttaaaatcacttcacacaacaa
 agtagcacaaagcagtttaacagaaagagcacatttgtggactccagatgggtcaggagatgagcaagg
 attggtatgtgactctgatgccccagcacagttacccccagcagcagagtgaaagaagtgtttgtgtggg
 ttttgttagtctggattcggatgtataaggtgtgccttgtactgtctgatttactacacagagaaacttt
 40 ttttttttttaagatatagactaaaatggacaattgtttacaaggcttaactaatattttgtctttt
 taaacttgaactttctgtataatagatagcttctttggattatgattttaagaaattattaattatgaa
 atgataggtaaggagaagctggattatctcctgttgagagcaagagattcgttttgacatagagtgaatg
 cattttccccctcctcctcctcctgctaccattatatttgggggttatgttttgcctctttaaagatagaaa
 tccagttctctaatttgggttttcttcttgggaaaccaaatacaaaatgaatcagtatcaattagggc
 45 ctggggtagagagacagaaaacttgagagaagagaagttagtgttccctctcttctctgttggtaggaa
 tcacctgaagacctagtcctcaatttaattgtgtgggtttttaaatttctcctagaatgaagtgcagtaa
 caatgagaaagaatacagcacaaccccttgaacaaaatgtatttagaaatatatttagttttatagcagaa
 gcgctcaattgtttgggttggaaagttaggggaaattgaagtgtagtctcagtgagaaatggctatgaa
 gcgtcatttcacattttaccccaactgacctgcatagccaggacacaaagttaaaacatttgtgagatagt
 50 gtggtaagtgtgactcgtgttaagtcaaaaggctataagaaacactgtgaaaggttcattatccatccat
 tgtgattcttccccacgtcttgcattgtattactggatccccacagtaatatagactgtgcattggtgtgt
 atatttcatttgcgatttccctgttaagatgagtttgtactcagaattgaccaattcaggagggtgaaaaat
 aaacagtggttctcttcttaccacaaagccactactgaccaaggtctcttcagtgcactcgctccctctc
 tggctaaggcatgcattagccactacacaagtcatttagtgaaagtggtcttttatgtcctcccagcagac
 55 agacatcaaggatgagtttaaccaggagactcctgtgactgtggagctctggaaggcttggtgggagt
 gaatttggccacaccttacaatttgtggcaggatccagaagagcctgtcttttatatccattccttgatg
 tcattggcctctcccacagatttcattacggtgccacgcagtcattggatctgggttagtcgggaaaaaaa
 aggagggaagacagcctggtaatgaataagatccttaccacagttttctcatgggaaatacataataaac
 cctttcatcttttttttttttctttaaagaattaaaactgggaaatagaaacatgaactgaaaagtcttgc
 60 aatgacaagaggtttcatgggtcttaaaaagatactttatatgggtgaagatgaaatcattcctaataata
 ccttttttttaaaaaaaacaatgtatattatgttccctgtgtgtgaatttaaaaaaaataacttta
 cttggatattcatgtaatatataaagggttgggtgaaatgaactttagttaggaaaaagctggcatcagct
 ttcattctgtgtaagttgacaccaatgtgtcataaatattctttattttgggaaattagtgtatttataaa
 aatttttaaaaaagaaaaagactactacaggttaagataatttttttaccctgtctttctccatattttaa
 gctatgtgattgaagtacctctgttcatagtttccctgtataaagttgggttaaaatttcattctgttaata
 gatcattaggttaataatgtatgggttttctattgggtttttgcagacagtagaggagatttggtaac
 aagggtctgttacacagtgataggttaataaataatgcaatttatcactccttttcatgttaataatt
 tgaggactggataaaagggtttcaagattaaaatttgatgttcaaacccttctgt

Figure 7: C terminus protein fragment of hPOSH (public gi:7959249; SEQ ID NO:7)

5 I SYVEFN SAAKQLIEWDKPPVPGVDAGECSSAAQSSSTAPKHSDTKKNTKKRHSFTSLTMANKSSQASON
RHSMEISPPVLISSSNPTAAARISELSGLSCSAPSQVHISTTGLIVTPPPSSPVTTGPSFTFPDVPYQA
ALGTLNPPPLPPPPLLAATVLASTPPGATAAAAAAGMPREMGSTDQIAHLRPQTRPSVYVAIYPYTPRK
EDELELRKGEMFLVFERCQDGWFKGTSMTSKIGVFPGNVAPVTRAVTNASQAKVPMSTAGQTSRGVTM
10 VSPSTAGGPAQKLQGNVAGSPSVVPAAVVSAAHIQTSPOAKVLLHMTGQMTVNQARNAVRTVAAHNQER
PTAAVTPIQVQNAAGLSPASVGLSHSLASPOAPLMPGSATHTAAISISRASAPLACAAAAPLTSPSIT
SASLEAEPGRIVTVLPGLPTSPDSASSACGNSSATKPKDKSKKEKKGLLKLKLSGASTKRKPRVSPASP
TLEVELGSAELPLQGAVGPELPPGGGHGRAGSCPVDGDGPVTTAVAGAALAQDAFHRKASSLDSAVPIAP
PPRQACSSLGPVLNESRPVVCERHRVVSYPPQSEAELELKEGDIVFVHKKREDGWFKGTLQRNGKTGLF
15 PGSPFVENI

ATTGTTTACAAGGCTTAACTAATTTATTTGCTTTTTTAAACTTGAACTTTTCGTATAATAGATACGTTCT
TTGGATTATGATTTTAAAGAAATTATTAATTTATGAAATGATAGGTAAGGAGAAGCTGGATTATCTCCTGT
TGAGAGCAAGAGATTTCGTTTTGACATAGAGTGAATGCATTTTCCCCTCTCCTCCTCCTGCTACCATAT
5 ATTTTGGGGTTATGTTTGGCTTCTTTAAGATAGAAATCCCAGTTCTCTAATTTGGTTTTCTCTTTGGGA
AACCAAACATACAAATGAATCAGTATCAATTAGGGCCTGGGGTAGAGAGACAGAACTTGAGAGAGAGA
AGTTAGTGATTCCTCTCTTTCTAGTTTGGTAGGAATCACCTGAAGACCTAGTCCTCAATTTAATTGTG
TGGGTTTTTAATTTTCTAGAAATGAAGTGAAGTGAACAATGAGAAAGATACAGCACAACTTGAACAA
10 AATGTATTTAGAAATATATTTAGTTTATAGCAGAAGCAGCTCAATTTGTTTGGTTGGAAAGTAGGGGAAA
TTGAAGTTGTAGTCACTGTCTGAGAATGGCTATGAAGCGTCATTTACATTTTACCCCAACTGACCTGCA
TGCCCAAGGACACAAGTAAACATTTGTGAGATAGTGGTGGTAAGTGATGCACTCGTGTAAAGTCAAAGGC
TATAAGAAACACTGTGAAAAGTTTATTTTATTCATCCATTGTGATTCTTTCCCAAGCTCTTGCATGTATTACT
GGATTTCCACAGTAATATAGACTGTGCATGGTGTGTATTTTCAATGCGATTTCCTGTAAAGATGAGTTT
GTACTCAGAATTGACCAATTGAGGAGGTGTAATAATAAACAGTGTCTCTTCTACCCCAAAGCCACTA
15 CTGACCAAGGTCTCTTTCAGTGCATCTGCTCCTCTCTGGCTAAGGCATGCATTAGCCACTACACAAGTCA
TTAGTGAAGTGGTCTTTTATGTCTCTCCAGCAGACAGACATCAAGGATGAGTTAACAGGAGACTACTC
CTGTGACTGTGGAGCTCTGGAAGGCTTGGTGGGAGTGAAATTTGCCACACCTTACAATTGTGGCAGGATC
CAGAAGAGCCTGTCTTTTATATCCATTCTTGTATGTCAATGGCCTCTCCACCGATTTTCAATTACGGTGC
CACGCAGTCATGGATCTGGGTAGTCCGGAACAAAGGAGGGAAGACAGCCTGGAATGAATAAGATCC
20 TTACCACAGTTTTCTCATGGGAAATACATAATAAACCTTTTCTCTTTTTTTTTTCTTTAAGAAATTAA
AACTGGGAAATAGAAACATGAACGTGAAAAGTCTTGCAATGACAAGAGGTTTCATGGTCTTAAAAAGATAC
TTTATATGGTTGAAGATGAATCATTCTCTAAATTAACCTTTTTTTTTAAAAAAAACAATGTATATTATGT
TCCTGTGTGTGAATTTAAAAAAAATACTTTACTTGGATATTCATGTAATATAAAGGTTTGGTG
25 AAATGAACTTTAGTTAGGAAAAAGCTGGCATCAGCTTTCATCTGTGTAAGTTGACACCAATGTGTATATA
TATTCCTTATTTTGGGAAATTAGTGATTTTATAAAAAATTTAAAAAAGAAAAAGACTACTACAGGTTAA
GATAATTTTTTACCTGTCTTTTCTCCATATTTAAGCTATGTGATTGAAGTACCTCTGTTTCATAGTTTC
CTGGTATAAAGTTGGTTAAATTTTCTGTTAATAGATCATTAGGTAATATAATGTATGGGTTTTCTAT
TGTTTTTTTGCAGACAGTAGAGGGAGATTTTGTAAACAAGGGCTTGTACACAGTGATATGGTAATGATAA
30 AATTGCAATTTATCACTCCTTTTCTGTTAATAATTTGAGGACTGGATAAAAGGTTTCAAGATTAAAAAT
TGATGTTCAACCTTTGT

Figure 9: Domain Analysis of Human POSH

Domain Name	begin	end	E-value
<u>RING</u>	12	52	1.06e-08
<u>SH3</u>	137	192	2.76e-19
<u>SH3</u>	199	258	4.84e-15
<u>low complexity</u>	366	384	-
<u>low complexity</u>	390	434	-
<u>SH3</u>	448	505	2.40e-19
<u>low complexity</u>	547	563	-
<u>low complexity</u>	652	668	-
<u>low complexity</u>	705	729	-
<u>SH3</u>	832	888	1.47e-14

Figure 10: Diagram of Human POSH Nucleic Acids

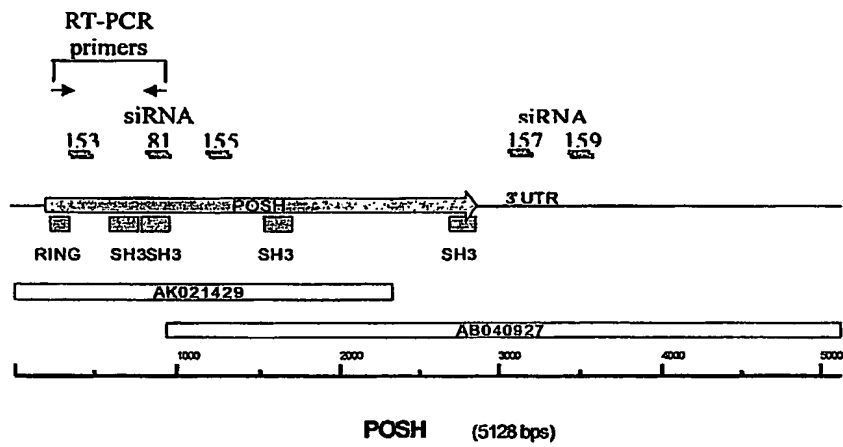


Figure 11: Reduction in Full Length POSH mRNA by siRNA Duplexes

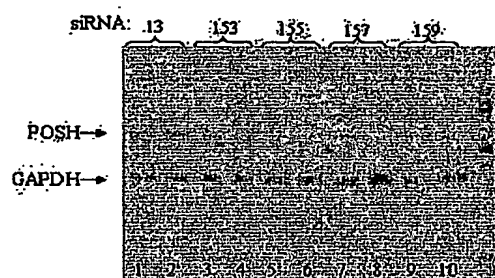


Figure 12: POSH Affects Release of VLP from Cells

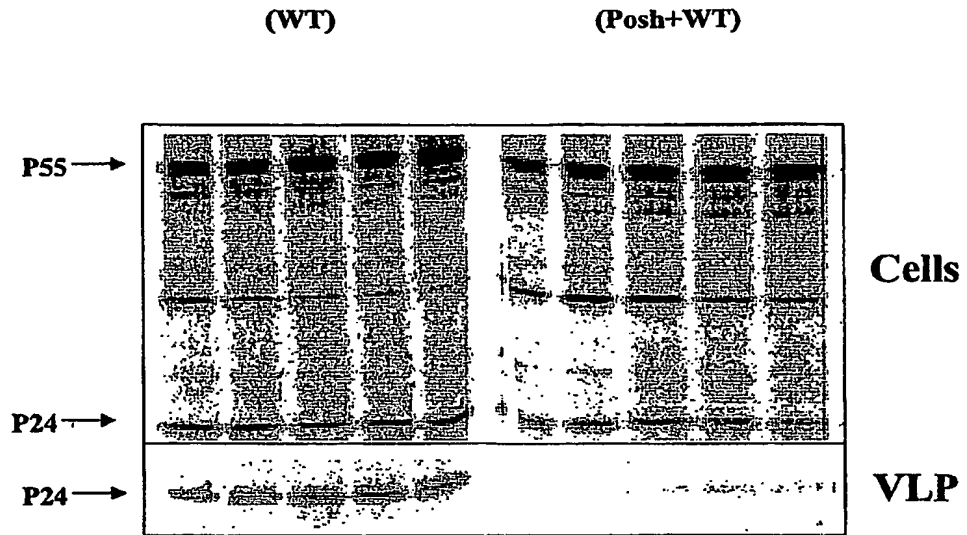


Figure 13: Release of VLP from Cells at Steady State

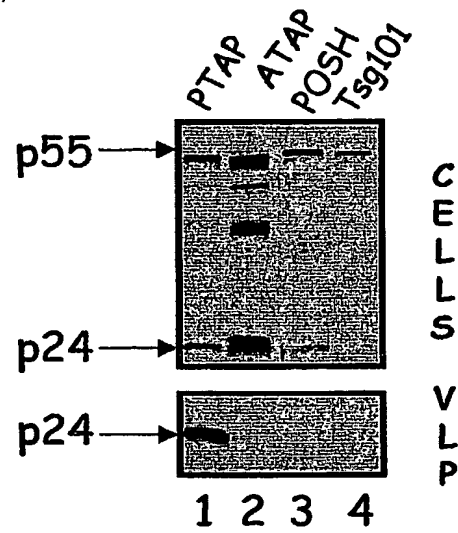


Figure 15: Mouse POSH Protein sequence (Public gi: 10946922; SEQ ID NO: 9)

5 MDESALLDLECPVCLERLDASAKVLPCQHTFCKRCLLGIVGSRNELRCPECRTLVGSGVDELPSNILLV
RLLDGIKQRPWKPGPGGGGGTTCNTLRAQGSTVVNCGSKDLQSSQCGQQPRVQAWSPPVRGIPQLPCAK
ALYNYEGKEPGDLKFSKGDITILRRQVDENWYHGEVSGVHGFFPTNFVQIIKPLPQPPQCKALYDFEVK
DKEADKDCLPFAKDDVLTVIIRVDENWAEGLADKIGIFPISYVEFNAAKQLIEWDKPPVPGVDTAECF
SATAQSTSASKHPDTKKNTRKRHSFTSLTMANKSSQGSQNRHSMEISPPVLISSSNPTAAARISELSGLS
CSAPSQVHISTTGLIVTPPPSSPVTTGPAFTFPSDVPYQAALGSMNPPLPPPPLLAATVLASTPSGATAA
10 VAAAAAAAAAAGMGRPVMGSSSEQIAHLRPQTRPSVYVAIYPYTPRKEDELELRKGEMFLVFERCQDGWY
KGTSMHTSKI GVFPGNYVAPVTRAVTNASQAKVSMSTAGQASRGVTMVSPSTAGGPTQKPKQNGVAGNPS
VVPTAVVSAAHIQTSPOAKVLLHMSGQMTVNQARNVARTVAHNSQERPTAAVTPIQVQNAACLGPASVGL
PHHSLASQPLPPMAGPAAHGAIVSISRTNAPMACAAGASLASPNMTSAMLETEPSGRTVTILPGLPTSPE
SAASACGNSSAGKPKDKSKKEKGLLKLKLSGASTKRKPRVSPPASPTLDVELGAGEAPLQGA VGPELPLG
15 GSHGRVGCPTDGDGPVAAAGTAALQDAFHRKTSSLDSAVPIAPPPRQACSSSLGPVMNEARPVVCERHRV
VVSYPQSEAELELKEGDIVFVHKKREDGWFKGTLQRNGKTGLFPGSFVENI

Figure 16: *Drosophila melanogaster* POSH mRNA sequence (public gi:17737480; SEQ ID NO:10)

```

5  CATTGTATCCGCTTGGCCACGAGCTTTGGCTGCACCTTGGCAAACCTAATAAATTAAACATTGAATCCTG
   CCTATTGCAACGATAATATAATCTGATTTAGTGCAATTAAGAACGACAAGTAGCGATTATAATAGTAGATT
   TTAGCATTGAGCTAAATTTATTTCCCAACCGCGTCTTGGGATTGCGTATGCGTGAGCCAGTACCTGCAT
   GTGTGTGTGTTTTGGAATGTGGCCCTGCACGAAATCAAATAGTGACCATCCTTGAGATTTTGCACTAG
10  GCAAGATGGACGAGCACACGTTAAACGACCTGTTGGAGTGCTCCGTGTGCTTGAGCGACTGGACACCAC
   ATCGAAGGTGCTGCCATGCCAGCACACCTTCTGCCGCAAATGCTTGACGACATTGTGGCCAGTCAGCAC
   AAGTTGCGATGCCCGGAGTGCCGCATCCTGGTCTCTTGCAAAATTGATGAGCTGCCCTCCAAACGTCTTGC
   TGATCGCAATCTTAGAAGGCATGAAACAAAATGCAGCAGCTGGCAAAGGAGAAGAAAAGGGAGAGGAGAC
   TGAACACACAGCCGAAAGGGCCAAACCTCAGCCGCCAGCGAATCAGTGGCCCCGCCCTGACAACCACTA
15  CTCCAGCTGCAGTCACATCAGCAATCTCATCAGCCGGCTCGTCACAAGCAACGTGATTTCTACTCCCCC
   ACGCTATGCCCCCTTTGACTTCGCCCTCCGGTGAAGCCACCGATCTAAAGTTCAAGAAAGGGGATCTGAT
   ACTGATCAAGCATCGCATCGACAACAACCTGTTTGTGGGTCAGCGAATGGTCAGGAGGGCACATTTCCC
   ATCAACTACGTCAGGTATCGGTTCCGCTGCCATGCCGAGTGCAATGGCCATGTATGACTTTAAGATGG
   GGGCCACGACGAGGAGGATGCCCTCGAATTTAAGAAAAGCACTGTAATACAGGTAATGCGCGAGTTGA
20  TCATAATTGGGCAGAAAGACGAATTGGCCAGACCATCGGAATCTTTCCAATAGCATTGCTTGAGCTGAAT
   GCAGCGGCCAAAAAGCTGTTGGACAGCGGGCTACACACCCATCCATTCTGCCATCCACCGAAGCAACAGG
   GGCAGCGGGCCCTTCTCCGGTTCCAGTTATGATCCACGGTGGTCAAGGAATCCAGTTCCGGATCCCTC
   CAATTCACGCGCGGAGCAGCAATTCAGCTCCACATCCAGCTCGAATAACTGCAGTCCGAATCACCAA
   ATCTCACTGCCGAATACCCCCAACATGTAGTAGCTTCCGGATCGGCGCTGTTTCGTTCCGTGACAAGG
25  GAGCAAGGAGAAACGCCACTCACTAAATGCTTTGCTGGGAGGAGGAGCTCCATTAACTCTGCTGCAGAC
   CAACCCCATTCGGCTGAAATTCCTAGCCTGCCCATGAACCTAAGCCGCTTGGAAAGTTCCAGCTCAACA
   GCTCTAAACCCACGTCAGCCCCACAGACATCGCGTGTACTTAAGACCACTGTTCCAGCAGCAGTGCAAC
   CGAATTTACCTGGGGATACTTAGCCCTGTTCCCATACAAACACGCCAAACGGATGAGCTGGAATTTAAA
   AAAGGGTTGTGTTACATTGTGACCGAACGATGTGTGACGGTGGTTCAAGGGAAAAAACTGTTGGAC
30  ATCACTGGAGTGTTCCCGGGCAACTACCTGACGCCCCCTGCGCGCCCGCAGCAGCAGTTAATGCATC
   AATGGAAATATGTTCCCCAAAATGCAGACGCCCAGATGGCAAGTACAGCAGCATCCAGTTGCACCAAGA
   TGTGCGACTCAACAACATGCTGTCCATGCAACCGCCTGATTTGCCACCTCGTCAGCAGCAGGCTACCGCC
   ACGACCAACAGTTGCTCTGTGTGGTGCAGAACAGTGGAGGCGCTGTTCCAGCAGAAAACTCGGAGCCCAAGC
   CTGAAACTGCCACAGCTTCGACTACGAGCAGCAGTTCTCTGGAGCAGTGGGACTTATGAGGAGATTAAAC
35  TCACATGAAAAACACGCTCCAAATCTCCGGGAGCGTCTTGCAGCAAGTTCGAAAGAAAGCTATTAGCACA
   AATGTGGAATTTACAACAAA CCCATCAGCTAAATTGCAATCCAGTACATGTAAGATCCGGCTCGTGCCCCA
   GTCAGCTGCAGCACAGTCAACCGCTCAATGAAACTCCAGCAGCCAAGACAGCGGCACAACAACAGCAGTT
   CCTACCCAAGCAGCTGCCCTTCCGCTTCTACGAACAGCGTTTCGTACGGATCGCAACCGGTGAAAGGAAGC
   AAGGAACGTCTCACTTGATTTGCGCGAGACAATCATTAGATGCAGCTACATTTCCGAGTATGTACAACA
40  ATGCCGCGTCCGCGCCGCCACCTACTACTTCCGTGGCCCCAGCTGTCTACGCGCGCGGTGAGCAACAGGT
   GATTCCTGGAGGTGGAGCGCAATCCAGTTGCATGCCAATATGATTATTGCCACCCAGCCATCGGAAGTCG
   CACAGCCTAGATGCGAGTCATGTGCTGAGTCCAGCAGCAATATGATCACGGAGGCGGCCATTAAAGCCA
   GCGCCACCACTAAGTCTCCTTACTGACAGAGGAAAGTCGATTCCGCTGCATGTGCGGTATCCACCAAA
   CAGTGACATTGAACTAGAGCTACATTGGGGCAGATTATCTACGTCAGCGGAAGCAGAAGAACGGCTGG
45  TATAAGGGCACCCATGCCCGTACCCACAAAACCGGGCTGTTCCCGCCTCCTTTGTTGAACCGGATTGTT
   AGGAAAGTTATGGTTCAAACTAGAATTTATTAAAGCGAAATTCAAAATTACTTGTCTAAAAGGATTCAATC
   GTCGGTCTATTCCGGCTTCCAAATACGCAATCTCATATTTCTCTTTCAAAAAAGAAACCGTTTGTACT
   CTTCCAATCGAATGGGCAGCTCGCGTTGTACTTTTTATACAATGCTTGATCAAAATAGGCTAGCCATG
   TAAGACTTAGGGAAACAGTTACTTAAGCCTTAGCGATTAGTTAGCTAGAGAAATAATCTAACCGATCCTTG
   TGCCCTCTACAAAGTTATTTGTAATATACGATACTCAGTAATAAAAAAAAAAAAAAAAAAAAAAAAAA

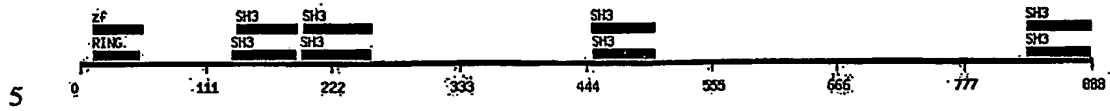
```

Figure 17: *Drosophila melanogaster* POSH protein sequence (public gi:17737481; SEQ ID NO:11)

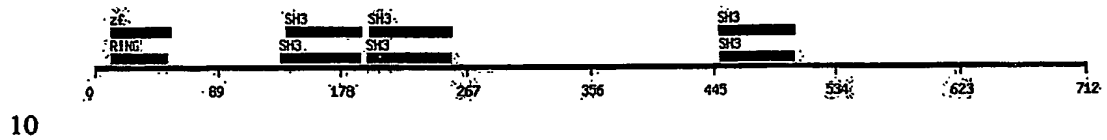
5 MDEHTLNDLLECSVCLERLDTTSKVLPCQHTFCRKCLQDIVASQHKLRCPCECRIIVSCKIDELPPNVLLM
RILEGMKQNAAGKGEEKGEETETQPERAKPQPPAESVAPPDNQLQLQSHQQSHQPARHKQRRFLLPHA
YALFDFASGEATDLKFKKGDLLIKHRIDNNWFVGOANGQEGTFPIINYVKVSVPLPMPQCIAMYDFKMGF
NDEEGCLEFKKSTVIQVMRRVDHNWAEGRIGQTIGIFPIAFVELNAAAKLLDSGLHHTPFCHPPKQQGQ
10 RALPPVPVIDPTVVTSSSSGSSNSTPGSSNSSSTSSSNCSFNHQISLPTPQHVVASGSASVRFRDKGA
KEKRHSINALLGGGAPLSLLQTNRHSAEILSLPHELRLVSSSTALKPTSAPQTSRVLKTTVQQQMOPN
LPWGYLALFPYKPRQTDELELKKGCYIIVTERCVDGWFKGKNWLDITGVFPNGNYLTPLRARDQQQLMHQW
KYVPQNADAQMAQVQHPVAPDVRLNNMLSMQPPDLPPRQQQATATTSCSVWSKPVEALFSRKSEPKPE
TATASTSSSSSGAVGLMRRLTHMKTRSKSPGASLQQVPKEAISTNVEFTTNPSAKLHPVHVRSGSCPSQ
15 LQHSQPLNETPAAKTAAQQQQFLPKQLPSASTNSVSYGSQSVKGSKERPHLI CARQSLDAATFRSMYNNA
ASPPPTTSVAPAVYAGGQQQVIPGGGAQSQLHANMIIAPSHRKSHSLDASHVLSPPSSNMITEAAIKASA
TTKSPYCTRESRFRCIVPYPPNSDIELELHLGDIIVYQRKQKNGWYKGTHTARTHTGLFPASFVEPDC

Figure 18: POSH Domain Analysis

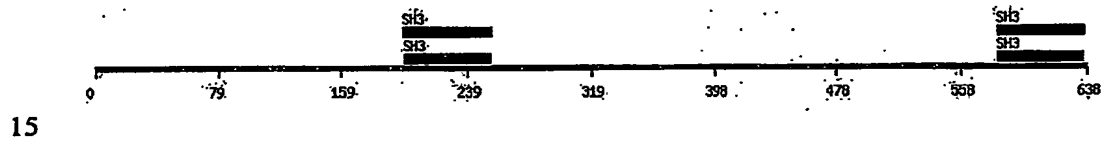
hPOSH protein sequence :



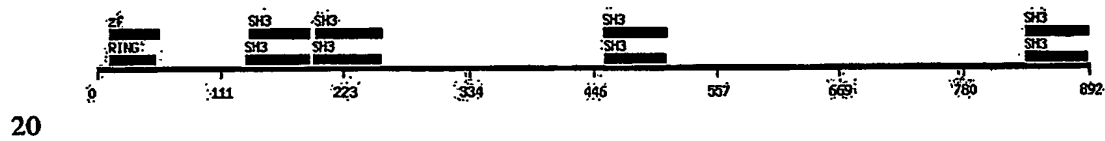
N terminus protein fragment of hPOSH (public gi:10432612):



C terminus protein fragment of hPOSH (public gi:7959249):



Mouse POSH Protein sequence (Public gi: 10946922):



Drosophila melanogaster POSH protein sequence (public gi:17737481)

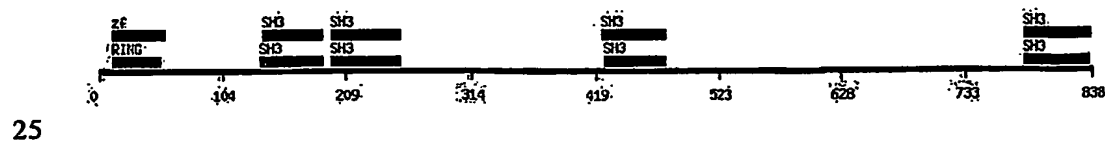


Figure 19

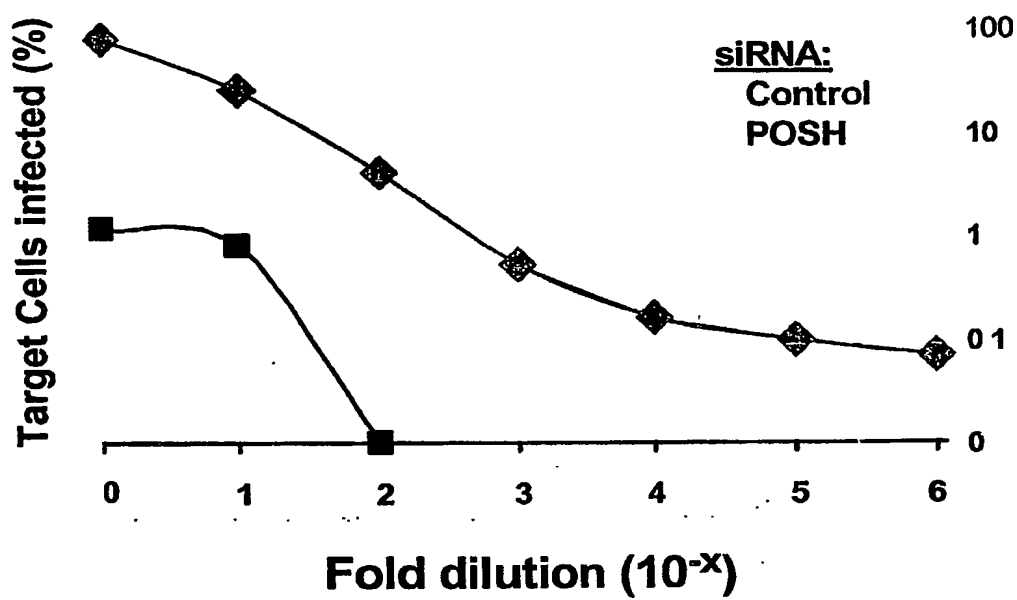


Figure 20.

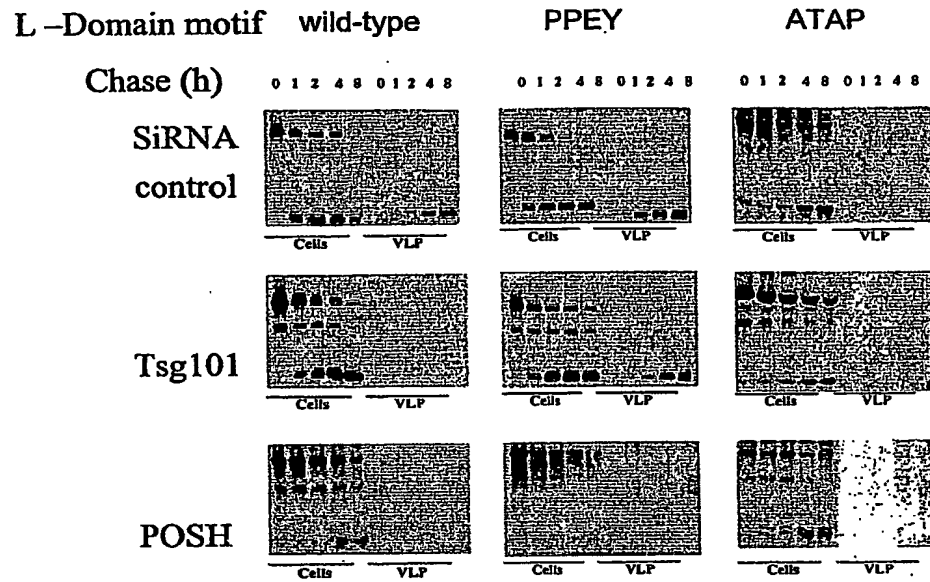


Figure 21

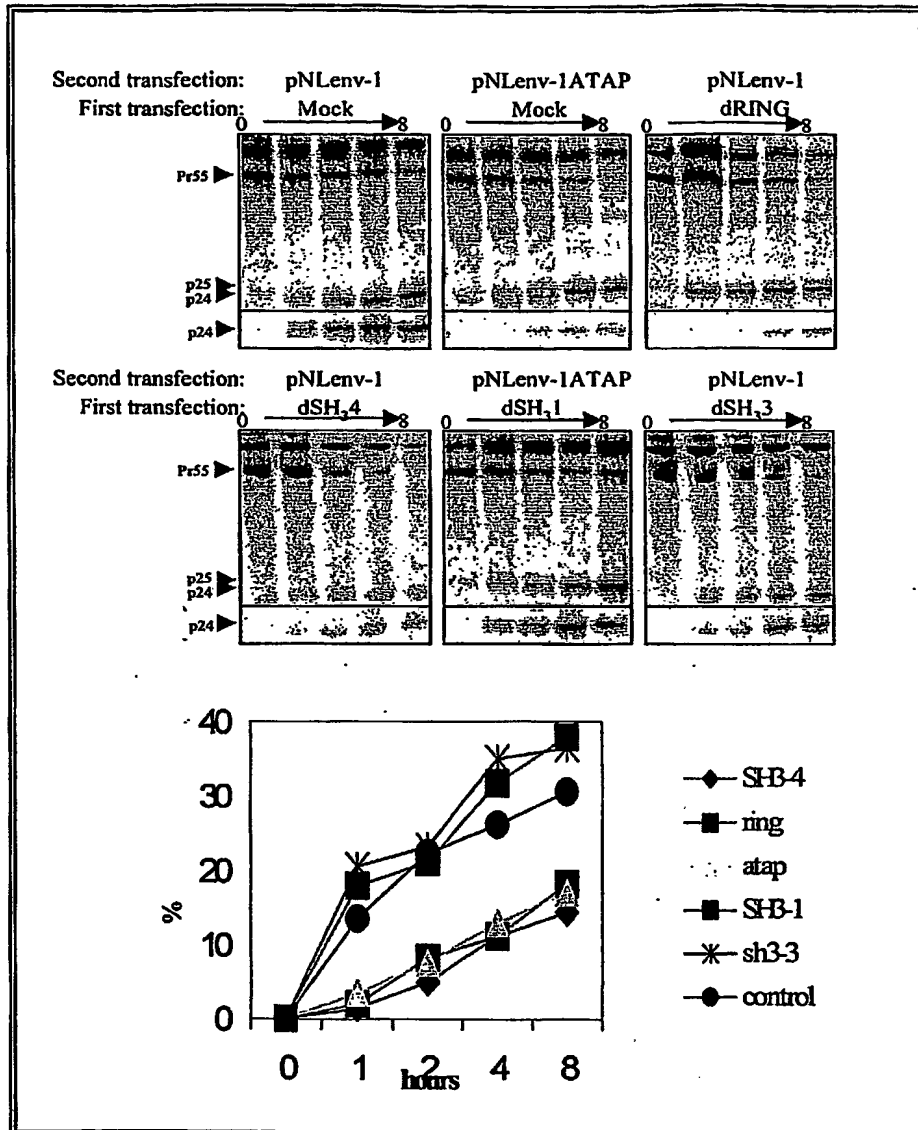


Figure 22: Human POSH has ubiquitin ligase activity

5

10

15

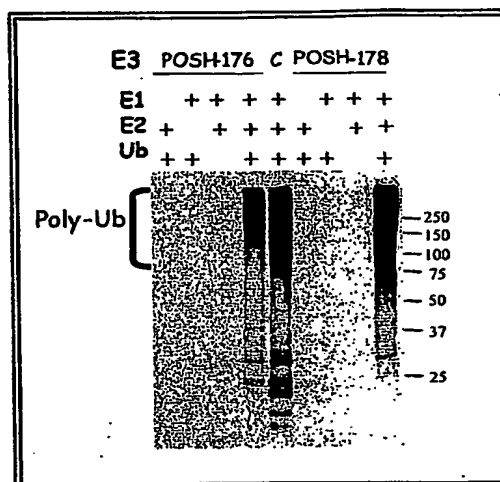


Figure 23

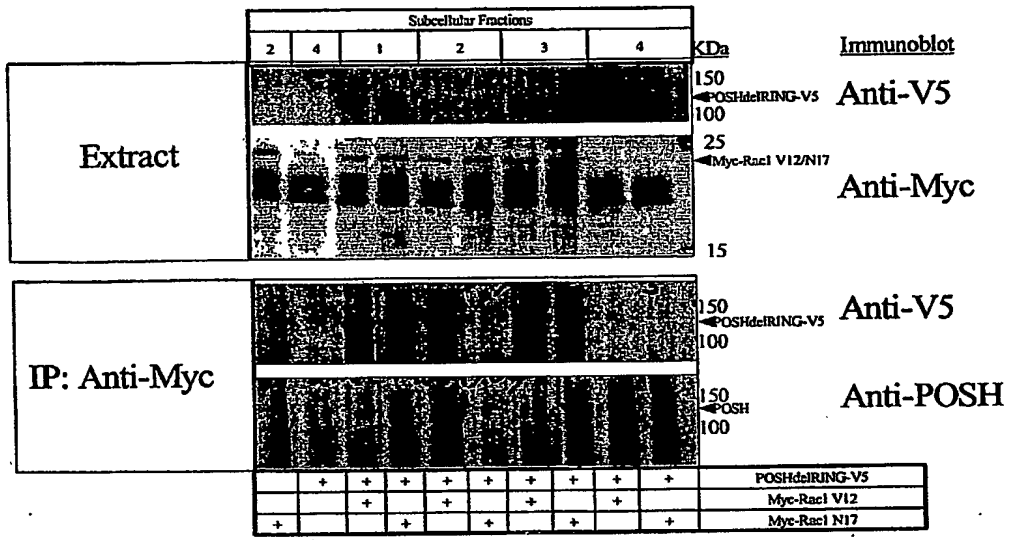
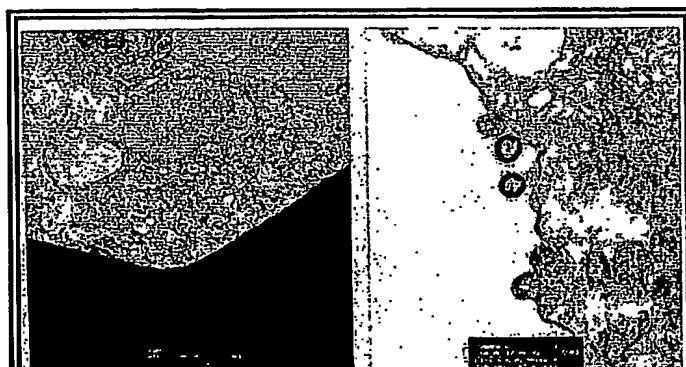


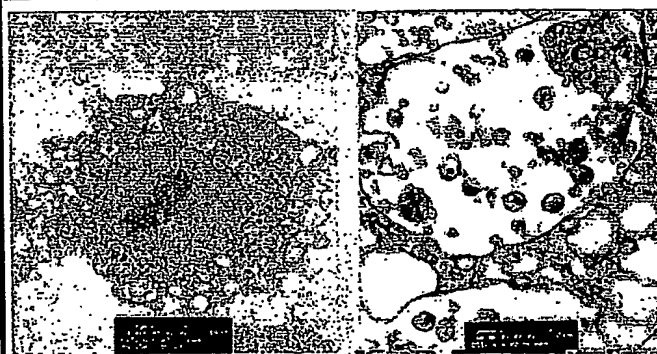
Figure 23

Figure 24. Knock-down of POSH entraps HIV virus particles in intracellular vesicles

SiRNA-Tsg101



SiRNA-POSH



Control

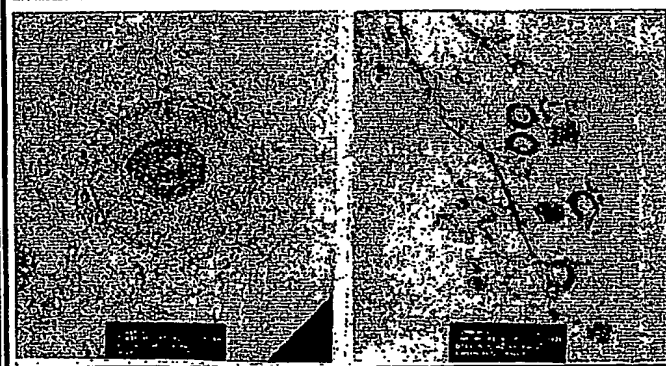


Figure 25. POSH is localized at two sites: one at the nuclei and one at the golgi (C).
After HIV transfection recruitment of POSH to the golgi is enhanced (D).

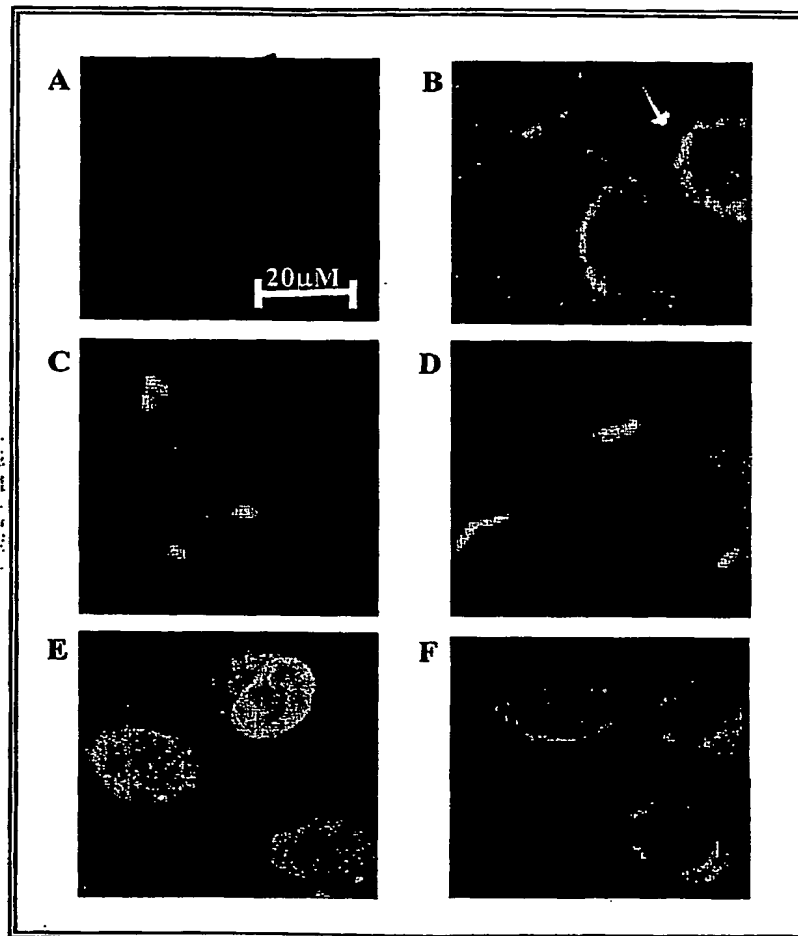
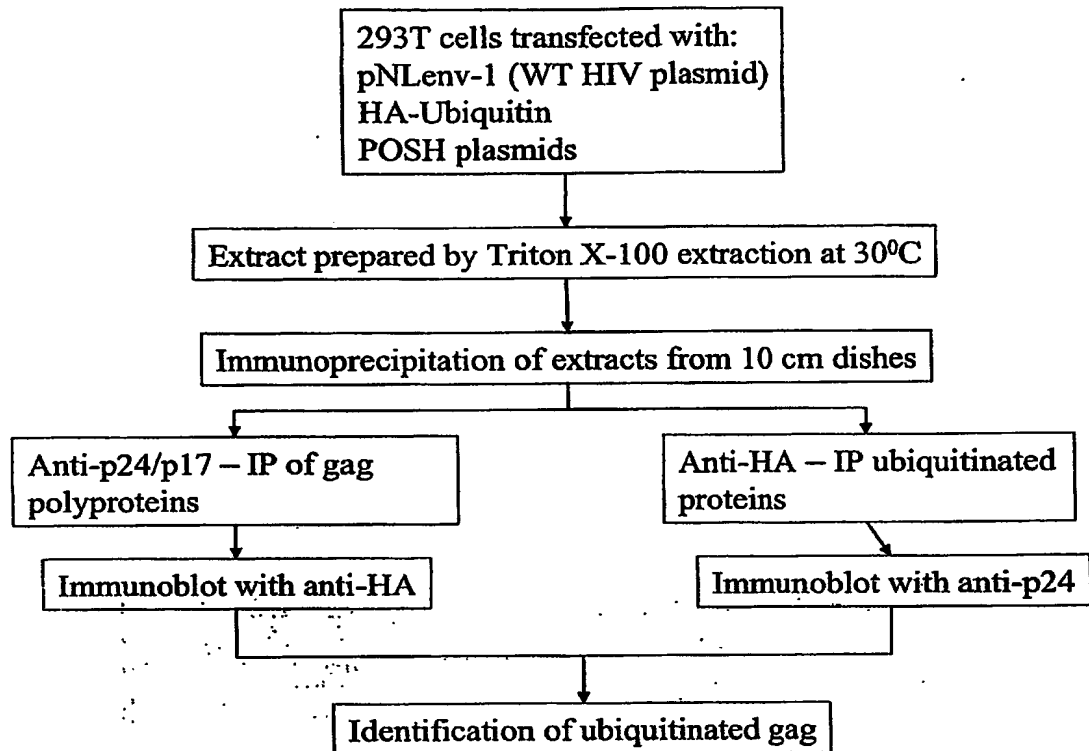
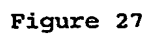


Figure 26

5





This Page is inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☒ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images
problems checked, please do not report the
problems to the IFW Image Problem Mailbox**